ANANTH KRISHNA PRASAD

Mail/Webpage:
ananth@cs.utah.edu / ananthkp.github.io

LinkedIn:
ananth-krishna-prasad-780b78116

AREAS OF INTEREST

Computer Architecture — Performance Architecture/Analysis — Processing-in-Memory — Hardware-Software Codesign — Memory System Design — Accelerating Big Data/Server-Scale Applications

RESEARCH PROJECTS

Silicon Performance Architect Intern, PhD May 2022 - August 2022
Reality Labs, Meta
Manager: Dr. Ahmad Samih

- **Memory Compression on the Edge** - Implemented and evaluated the performance of an on-chip IP near memory controller for Memory Compression. The proposed controller is capable of intelligently applying base-delta and frequent-pattern compression techniques depending on data characteristics. The memory controller was implemented using gem5 and we observed around 20% off-chip bandwidth reduction across a wide variety of benchmarks from SPEC, NAS and CRONO benchmark suites

Graduate Research Assistant August 2018 - **Present**
Kalhert School of Computing, University of Utah, UT
Advisors: Prof. Rajeev Balasubramonian & Dr. Mahdi Nazm Bojnordi

- **Billion-Scale Approximate Nearest Neighbor Search (ANNS)** - We propose a novel hardware-friendly algorithm for distance-based indexing of data in-memory through novel distance reformulation and pruning mechanism. We design a memory system capable of performing the required operations using a combination of PIM techniques. We demonstrate how such an approach could overcome the pitfalls associated with the *the curse of dimensionality* while not being bandwidth-bound. On a variety of vector datasets, we observe 52.7x/97x throughput/latency improvements over a server-class system, alog with a 4.15x energy reduction. **Under review** August 2021 - **Present**

- **LNS for DNN inference** - Logarithmic Number System for end-to-end DNN execution using only LUTs, bit-shifts and adders. The proposed approach shows around 95.2% acccuracy on VGG16 with Imagenet April 2021 - December 2021

- **Error model for Analog Crossbar Accelerators** - Modelled device non-idealities incurred as part of MAC operations performed on Crossbars. Devised a model integrated with NVSIM to identify inaccuracies in computations, parameterized by cell variance. Was used as part of XCRYPT publised in Transactions on Computing December 2020 - December 2021

- **Memristive Ranking In Memory** - Hardware/Software mechanism for large-scale data ranking in-memory. Identified bandwidth bottleneck issues with sorting kernels, and proposed viable hardware/software mechanism for performing large-scale data ranking in-memory with a bandwidth complexity of O(1), by reformulating sorting operations as bit-level in-situ operations. Achieved 12.4 - 50.7x throughput gains for high-performance parallel sorting kernels and 2.3 - 43.6x improvements in a set of database applications, with 90% energy reduction. September 2019 - December 2020

- **High Bandwidth Cross Caching** - Polymorphic Memristor-based memory with support for both caching and hashing. Developed a novel reconfigurable memristor based memory with high bandwidth efficiency, with capability of large scale parallel search. Demonstrated cache/scratchpad reconfigurability and achieved 50% and 12x improvement over state-of-the-art High Bandwidth memory, over Cache and Hash Table/Stringmatch applications respectively. December 2018 - May 2021

Research Assistant
IISc Bangalore, India July 2017 - June 2018
Advisor: Prof. S.K Nandy

- **WCET guarantees for CGRA** - Implemented and validated Worst Case Execution Time (WCET) analysis over the REDEFINE CGRA for validation of safety-critical application execution.

Course Projects

- **Energy Efficient-Transfer of Data in CNN accelerators by minimizing wire bit-flips**, Advanced Computer Architecture (Spring 2020)
- **Exploring Federated Learning**, Neuromorphic Architectures (Fall 2019)

PUBLICATIONS

- **Memristive Data Ranking** - **Ananth Krishna Prasad**, Morteza Rezaalipour, Masoud Dehyadegari, Mahdi Nazm Bojnordi, International Symposium on *High Performance Computer Architecture* 2021.
- **Monarch: A Durable Polymorphic Memory For Data Intensive Applications** - **Ananth Krishna Prasad**, Mahdi Nazm Bojnordi, IEEE *Transactions on Computers* 2022.
- **Enabling Distance-based Addressing in Non-Volatile Memory systems** - **Ananth Krishna Prasad**, Mahdi Nazm Bojnordi, Rajeev Balasubramonian, to appear in *Non-Volatile Memories Workshop* 2023.
- **XCRYPT: Accelerating Lattice Based Cryptography with Memristor Crossbar Arrays** - Sarabjeet Singh, Xiong Fan, **Ananth Krishna Prasad**, Lin Jia, Anirban Nag, Rajeev Balasubramonian et al., *IEEE MICRO* 2023

TECHNICAL BLOGPOSTS/POSTER

- **Nanoscale Optoelectronic AI Processing**, Ananth Krishna Prasad, Mahdi Nazm Bojnordi May 7, 2021
- **A Case for Optical Deep Neural Networks**, Ananth Krishna Prasad and Mahdi Nazm Bojnordi Oct 2, 2020
- **A Case for the Scope of Reconfigurable Transistors in Computer Architecture**, Ananth Krishna Prasad, Pierre-Emmanuel Gaillardon, Mahdi Nazm Bojnordi May 16, 2019
- High Bandwidth Cross Caching, presented at DAC 2020

EDUCATION

Doctor of Philosophy, Computer Science August 2018 - *Present*
University of Utah GPA 3.904

Bachelors in Technology, Electronics and Communication Engineering August 2013 - May 2017
Birla Institute of Technology and Science, Pilani (Hyderabad Campus), India
GPA - 8.35 out of 10