SASINDU WIJERATNE CURRICULUM VITAE

Email: kangaram@usc.edu **Phone Number**: +1 (213) 476-3375
Website: kangaram.wixsite.com/personal

RESEARCH INTERESTS

- Parallel and Distributed Computing
- Hardware-Software Co-design
- Tensor Decomposition
- Graph Algorithms
- Memory Optimizations
- FPGA Accelerators

EDUCATION

Doctor of Philosophy in **Computer Engineering** *Aug, 2019 - Current*
University of Southern California, Los Angeles, CA **CGPA: 3.94/4.00**

B.Sc. Eng (Hons.) in **Electronic & Telecommunication Engineering** *Jan, 2014 - Jan, 2018*
Univeristy of Moratuwa, Sri Lanka **CGPA: 3.75/4.20**

SELECTED PUBLICATIONS

- **Wijeratne, S.**, Kannan, R., Prasanna, V., 2021, Sep. Reconfigurable Low-latency Memory System for Sparse Matricized Tensor Times Khatri-Rao Product on FPGA. In *Proceedings of the 2021 IEEE High Performance Extreme Computing Conference*, Boston, USA
- **Wijeratne, S.**, Pattnaik, S., Chen, Z., Kannan, R., Prasanna, V., 2021, Aug. Programmable FPGA-based Memory Controller. In *Proceedings of the 2021 IEEE 28th Hot Interconnects Symposium*, Online
- Zhang, R., **Wijeratne, S.**, Yang, Y., Kuppannagari, S., Prasanna, V., 2020, July. A High Throughput Parallel Hash Table on FPGA using XOR-based Memory In *2020 IEEE High Performance Extreme Computing Conference*, Massachusetts, USA
- **Wijeratne, S.**, Ekanayake, A., Jayaweera, S., Ravishan, D., Pasqual, A., 2019, Feb. Scalable High Performance SDN Switch Architecture on FPGA for Core Networks. In *Proceedings of the 2019 ACM/SIGDA International Symposium on Field-Programmable Gate Arrays*, New York, NY, USA
- **Wijeratne, S.**, Jayaweera, S., Dananjaya, M., Pasqual, A., 2018, July. Reconfigurable Co-Processor Architecture with Limited Numerical Precision to Accelerate Deep Convolutional Neural Networks. In *proceedings of the 2018 IEEE 29th International Conference on Application-specific Systems, Architectures and Processors (ASAP) 2018*, Politecnico di Milano, Milan, Italy
- Senanayake, S., Liyanage, N., **Wijeratne, S.**, Atapattu, S., Athukorala, K., Tharaka, P., Karunaratne, G., Senarath, R., Perera, I., Ekanayake, A., Pasqual, A., 2017, July. High performance hardware architectures for Intra Block Copy and Palette Coding for HEVC Screen Content Coding Extension. In *IEEE 28th International Conference on Application-specific Systems, Architectures and Processors (ASAP) 2017*, Seattle, WA, USA

INDUSTRY EXPERIENCE

Engineer, Wave Computing *Feb, 2018 - July, 2019*
Wave Computing, Santa Clara, USA

Wave Computing is a Silicon-Valley based start-up developing a massively parallel data-flow architecture called Wave DataFlow Processing Unit(DPU) for high-speed deep learning solutions. I Engage in developing and implementing Machine Learning applications on Data Processing Unit (DPU) which is a novel Coarse-Grained re-configurable architecture using in-house programming tools.

Intern, Video Processing Division *Aug, 2016 - Dec, 2016*
ParaQum Technologies (Pvt) Ltd, Dehiwala, Sri Lanka

Worked as a visiting undergraduate researcher in the Video Encoding Division of ParaQum Technologies focusing on FPGA implementation of High Efficiency Video Coding Encoder and its Screen Content Coding extension, including designing and implementing a hardware friendly Palette Encoding Scheme and integrating it to their existing all-intra HD HEVC encoder.

Student Trainee, Research Institute for Nanodevice and Bio Systems *July, 2017 - Aug, 2016*
Research Institute for Nanodevice and Bio Systems, Hiroshima University, Hiroshima, Japan

Completed a Digital IC design fabrication training program on CMOS technology.

AWARDS AND ACHIEVEMENTS

- Represented University of Moratuwa, Sri Lanka in IEEE Signal Processing Cup 2017, was among final 20.
- All Sri Lanka Rank 81 (out of ∼60,000) in University Entrance Examination.
- All Sri Lanka Rank 16 (out of ∼350,000) in General Ordinary Level Examination.
- Mahapola merit scholarship to pursue undergraduate studies by the government of Examination.

RESEARCH EXPERIENCE

Shared Memory Controller for Hardware Accelerators *2020 Jan - present*
University of Southern California, Supervised by Prof. Viktor Prasanna
With the rise of Big Data, there has been a significant effort in increasing compute power through heterogeneous architectures. As a result, many applications are memory-bound, i.e., they are bottlenecked by the movement of data from the main memory to compute units. There is a massive trend towards using heterogeneous hardware architectures (ex: CPU & GPU, CPU & TPU) to accelerate computationally intensive algorithms (Eg: Machine Learning). Conventional memory access techniques have become obsolete with the current heterogeneous systems as their batch computational time is significantly low, and data sharing among processing units are more frequent than ever before. There are novel massively parallel platforms that required recurring memory accesses with huge bandwidths. Conventional Memory Controllers are unable to handle such data communication between processing elements and main memory. Therefore, in this project, we are proposing a Memory Controller that can handle such data traffic with intelligent data routing/mapping techniques.

Offloading Distributed Graph Processing to the Public Cloud? Evaluating the Impact of Communication Slowdowns *2019/20*
University of Southern California, Supervised by Prof. Viktor Prasanna
Extreme scale graph analytics is imperative for several real-world applications with the underlying graph structure containing millions or billions of vertices and edges. As such huge graphs cannot fit into the memory of a single computer, distributed processing of the graph is required. As a local dedicated cluster of machines with a low latency network can be prohibitively expensive, public clouds are becoming a viable alternative due to their flexibility and elasticity under a pay-as-you-go model. In this work, we develop performance models for three of the most popular communication patterns used by existing distributed graph processing frameworks.

Re-configurable co-processor architecture with limited numerical precision to accelerate deep convolutional neural networks *2016/18*
University of Moratuwa, Supervised by Dr. Ajith Pasqual
Designed a scalable limited precision model independent reconfigurable co-processor with a throughput of up to 226.2 GOp/s on Xilinx Virtex 7 FPGA. This architecture consists of parallel Multiply and Accumulate Units(MAC) with caching techniques and interconnection networks to exploit maximum data parallelism.

Software Defined Networking(SDN) Switch for Core-Networks on FPGA *2017/18*
University of Moratuwa, Supervised by Dr. Ajith Pasqual
Implemented a scalable high-performance Software Defined Networking(SDN) switch fabric for the core-network environment. The implemented FPGA-based switch which is fully compliant with OpenFlow; the pioneering protocol for the southbound interface of SDN. The of an OpenFlow Southbound agent which can process OpenFlow packets at a rate of 10Gbps. The speed of overall architecture scales up to 400Gbps on a Xilinx Virtex-7.

HEVC Encoder - SCC Extension *2016/17*
ParaQum Technologies, Supervised by Dr. Ajith Pasqual
Introduced a resource efficient real-time hardware designs for the major SCC tools including Intra Block Copy, and Palette Coding. Moreover, a hardware-friendly palette indices coding scheme is suggested for Palette

Coding with significant hardware utilization. These designs are targeted to achieve the throughput necessary for a 1080p 30 frames/s encoder and synthesized for a Virtex-7 VC707 evaluation platform.

COURSES & TECHNICAL SKILLS

- **Software & Programming**
  Verilog, VHDL, OpenMP, MPI, CUDA, C/C++, Python, Matlab, ModelSim
- **Coursework**
  Analysis of Algorithms, Machine Learning, Computer Systems Architecture, Parallel and Distributed Computing, Computer Systems Organization,

MISCELLANEOUS

- Teaching Assistant for Accelerated Computing using FPGAs (EE 599), Introduction to Computer Systems (CSCI 356) and Introduction to Digital Circuits (EE 354L) at University of Southern California.