Hyoukjun Kwon

Ph.D. Student
Klaus Advanced Computing Building 3305, 266 Ferst Dr. NW, Atlanta, GA 30332
(+1) 404-539-4457 | hyoukjun@gatech.edu | www.hyoukjunkwon.com | hyoukjun-kwon

Research Interest

Deep Neural Network (DNN) accelerators
Interconnection network (NoCs)
Reconfigurable architecture
Computer architecture
Machine learning
Design automation
Hardware security

Education

Georgia Institute of Technology
Ph.D. student in Computer Science — *Aug. 2015 - Current*
Advisor: Dr. Tushar Krishna

SNU (Seoul National University)
B.S. in CSE (Computer Science and Engineering)
B.S. in EMS (Environmental Material Science) — *Mar. 2007 - Feb. 2015*
Advisor: Dr. Jihong Kim (CSE) and Dr. Junjae Lee (EMS)

Professional Experience

Facebook, Menlo Park, CA
Research Intern at Reality Lab — *May 2018 - July 2018*
(Accepted Offer)

NVIDIA, Westford, MA
Research Intern at architecture research group (ARG) — *May 2018 - Aug. 2018*
Manger: Dr. Steve Keckler, Mentor: Dr. Michael Pellauer

NVIDIA, Westford, MA
Research Intern at architecture research group (ARG) — *May 2017 - Aug. 2017*
Manger: Dr. Steve Keckler, Mentor: Dr. Michael Pellauer

Georgia Institute of Technology, Atlanta, GA
Graduate Research Assistant — *Aug. 2015 - present*

Talks in Industry

An Open Source Framework for Exploring Dataflow and Generating DNN Accelerators Supporting Flexible Dataflow
IBM Research, Yorktown Heights, New York — *Nov. 2018*

Communication-driven Approach to Design DNN Accelerators
Western Digital, WebEx Talk *Sep. 2018*

Analyzing Dataflows in Accelerators
NVIDIA, Westford, Massachusetts *Aug. 2018*

Optimizing Networks-On-Chip for Deep Learning Accelerators using Micro-switches
NVIDIA, Westford, Massachusetts *Aug. 2017*

Automatic generation of low-latency networks-on-chip
Bluespec Inc., Framingham, Massachusetts *Jul. 2017*

Talks in Academia

Modeling and Analyzing Dataflows in DNN Accelerators
Tokyo City University, 2018, Tokyo, Japan *Dec. 2018*

Enabling Rapid Design Space Exploration and Prototyping of DNN Accelerators
Massachusetts Institute of Technology (MIT), 2018, Cambridge, Massachusetts *Jul. 2018*

A Communication-driven Approach to Designing Flexible DNN Accelerators
CMU, Pittsburgh, Pennsylvania *May. 2018*

[Teaching] Designing CNN Accelerators using Bluespec System Verilog
Seoul National University (SNU), Seoul, Korea *Dec. 2017*
A three-day lecture for undergraduate students
(`https://github.com/hyoukjun/DesignCNNAccelerators`)

Light-weight and High-performance NoC for DNN accelerators
Konkuk University, Seoul, Korea *Oct. 2017*

Talks in Conferences/Workshops

[Tutorial] MAERI: Enabling Rapid Design Space Exploration and Prototyping of DNN Accelerators
ISCA 2018, Los Angeles, California *Mar. 2018*
`http://synergy.ece.gatech.edu/tools/maeri/maeri_tutorial_isca2018/`

MAERI: Enabling Flexible Dataflow Mapping over DNN Accelerators via Reconfigurable Interconnects
ASPLOS 2018, Williamsburg, Virginia *Mar. 2018*

[Demo] MAESTRO: An Open-source Infrastructure for Modeling Dataflows within Deep Learning Accelerators
Corgarch 2018 (colocated with ASPLOS 2018), Williamsburg, Virginia *Mar. 2018*

[Demo] OpenSMART: An Opensource Single-cycle Multi-hop NoC Generator
SC17 OpenSuCo 2017, Denver, Colorado *Nov. 2017*

Rethinking NoCs for Spatial Neural Network Accelerators
NOCS 2017, Seoul, Korea *Oct. 2017*

Adaptive Manycore Architectures for Big Data Computing
NOCS 2017, Seoul, Korea *Oct. 2017*

Proving Flow Security of Sequential Logic via Automatically-Synthesized Relational Invariants
CSF 2017, Santa Barbara, California *Aug. 2017*

OpenSMART: Single-cycle Multi-hop NoC Generator in BSV and Chisel
ISPASS 2017, Santa Rosa, California *May. 2017*

Publications

Zhongyuan Zhao, **Hyoukjun Kwon**, Sachit Kuhar, Weiguang Sheng , Zhigang Mao, and Tushar Krishna, *"mRNA: Enabling Efficient Mapping Space Exploration on a Reconfigurable Neural Accelerator"*, *International Symposium on Performance Analysis of Systems and Software (**ISPASS**), 2019 (To appear in the conference)*

Hyoukjun Kwon, Ananda Smajdar, and Tushar Krishna, *"A Communication-driven Approach for Designing Flexible DNN Accelerators"*, *IEEE Micro Special Issue on Hardware Acceleration **IEEE Micro**, 2018 (to appear in the journal)*

Brian Lebiednik, Sergi Abadal, **Hyoukjun Kwon** and Tushar Krishna, *"Architecting a Secure Wireless Network-on-Chip"*, *International Symposium on Network-on-Chips (**NOCS**), 2018*

Hyoukjun Kwon, Ananda Samajdar, and Tushar Krishna, *" MAERI: Enabling Flexible Dataflow Mapping over DNN Accelerators via Reconfigurable Interconnects"*, *International Conference on Architectural Support for Programming Languages and Operating Systems (**ASPLOS**), 2018*

Hyoukjun Kwon, Ananda Samajdar, and Tushar Krishna, *" MAERI: Enabling Flexible Dataflow Mapping over DNN Accelerators via Reconfigurable Interconnects"*, *The Inaugural Sysml Conference (**Sysml**), 2018*

Brian Lebiednik, Sergi Abadal, **Hyoukjun Kwon**, and Tushar Krishna, *"Spoofing Prevention via RF Power Profiling in Wireless Network-on-Chip"*, *International Workshop on Advanced Interconnect Solutions and Technologies for Emerging Computing Systems (**AISTECS**), 2018*

Hyoukjun Kwon, Ananda Samajdar, and Tushar Krishna, *"Rethinking NoCs for Spatial Neural Network Accelerators"*, *International Symposium on Networks-on-Chips (**NOCS**), 2017*

Janardhan Rao Doppa, Ryan Gary Kim, Mihailo Isakov, Michel A. Kinsy, **Hyoukjun Kwon**, and Tushar Krishna, *"Adaptive Manycore Architectures for Big Data Computing"*, *International Symposium on Networks-on-Chips (**NOCS**)*, *2017*

Hyoukjun Kwon, William Harris, and Hadi Esmaeilzadeh, *"Proving Flow Security of Sequential Logic via Automatically Synthesized Relational Invariants"*, *Computer Security Foundations (**CSF**)*, *2017*

Hyoukjun Kwon and Tushar Krishna, *"OpenSMART: Single-Cycle Multi-hop NoC Generator in BSV and Chisel"*, *International Symposium on Performance Analysis of Systems and Software (**ISPASS**)*, *2017*

Hyoukjun Kwon, Dohyun Kim, Jisung Park, and Jihong Kim, *"Improving the Lifetime of NAND Flash-based Storages Using MADE (Minhash-Assisted Delta-compression Engine)"*, *Korean Institute of Information Science and Engineers Annual Conference (KIISE)*, *2014*

Preprint

Hyoukjun Kwon, Michael Pellauer, and Tushar Krishna, *"An Analytic Model for Cost-Benefit Analysis of Dataflows in DNN Accelerators"*, *Arxiv Preprint: arXiv:1805.02566*, *2018*

Book Chapter

Tushar Krishna, **Hyoukjun Kwon**, Ananda Samajdar, Michael Pellauer, and Angshuman Parashar, *"Synthesis lecture on computer architecture: Data Orchestration in DNN Accelerators"*, *Morgan & Claypool*, *Planned to be published in 2019*

Research Projects

Exploiting data and compute reuse in DNN accelerators

Georgia Institute of Technology *Oct. 2017 - PRESENT*

- Developing a DNN accelerator-framework co-design method to exploit both of data and compute reuse in DNN accelerators

Modeling performance and energy cost/benefits of neural network mapping on accelerators

Georgia Institute of Technology *Oct. 2017 - PRESENT*

- Developed an analytic model for CNN problem mapping on accelerator architectures
- Submitted a paper as the first author. Arxiv version is available: `https://arxiv.org/abs/1805.02566`

Fine-grained programmability support in DNN accelerators

Georgia Institute of Technology *Oct. 2017 - PRESENT*

- Developing a fine-grained ISA (Instruction Set Architecture for DNN accelerators to support dataflow programming
- Submitted a paper as the second author

DARPA - CHIPS project (Communication architecture support for highly modular chip design)

Georgia Institute of Technology *Sep. 2017 - PRESENT*

- Working on architecture level exploration

Enabling flexible dataflow in DNN accelerators

NVIDIA and Georgia Institute of Technology *May. 2017 - PRESENT*

- Developed a light-weight and reconfigurable NoC architecture to support flexible dataflow in DNN acclerators
- Developed a dataflow evaluation framework that can be pluged into DNN accelerator RTL
- Submitted a paper as the first author

A communication-aware flexible dataflow DNN accelerator (MAERI)

Georgia Institute of Technology *Jan. 2017 - Jan. 2018*

- Developed a new DNN accelerator microarchitecture for CNN and RNN that offloads computation into network side and maximize the utilization of computation units.
- Implemented RTL using Bluespec System Verilog.
- Published a paper as the first author in ASPLOS 2018 and Sysml.

Secure wireless NoC system

Georgia Institute of Technology *Aug. 2016 - Jan. 2018*

- Worked on a method that prevents hardware Trojan that monopolies wireless network channel bandwidth and slows down other cores.
- Published a paper as a co-author in AISTECS 2018

Scalable NoC design for DNN accelerators

Georgia Institute of Technology *Aug. 2016 - May. 2017*

- Developed a light-weight and low-latency interconnection network for CNN accelerators.
- Implemented distributed architecture of network fabric that maximizes PE utilization.
- Implemented RTL using Bluespec System Verilog.
- Published a paper as the first author in NOCS 2017.

Advanced low-latency network-on-chip generator

Georgia Institute of Technology *Nov. 2015 - Oct. 2016*

- Developed a network-on-chip RTL generator that provides state-of-the-art mesh and single-cycle-multi-hop (SMART) network.
- Verified the design comparing RTL simulation results with those of Garnet under synthetic workloads and synthesized the design for ASIC and FPGA
- Provided the source code as opensource; now it is used by Bluespec Inc. IIT Madras for RISC-V Shakti Project, and Intrinsix Corp.
- Published a paper as the first author in ISPASS 2017.

Information flow-secure FPGA accelerators

Georgia Institute of Technology *Oct. 2015 - Aug. 2016*

- Developed an information flow policy language and a policy checker using a state-of-the-art formal verification method.
- Successfully identified insecure open-source hardware designs within several minutes on a laptop.
- Published a paper as the first author in CSF 2017.

Endurance-aware flash transition layer design

Seoul National University *Dec. 2014 - Sept. 2015*

- Developed an FTL (flash-transition-layer) that delta-compresses write data toward flash memory and efficiently manages the mapping information of delta-compression
- Reduced significant write traffic toward flash memory, which enhances the lifetime of flash memories.
- Published a paper as the first author in KIISE 2014.

Teaching

Teaching assistant in advanced computer architecture course (CS6290/ECE6100)

Georgia Institute of Technology *Sep. 2018 - Current*

- Helped students in recitation session and office hour
- Q&A via Piazza forum

Special lecture: Designing a CNN accelerator Using Bluespec System Verilog

Seoul National University *Dec. 2017*

- Taught a 3-day course about CNN accelerator design for CS and ECE undergraduate students.
- Developed lecture materials and lab code, which lets students implement a full CNN accelerator.

Teaching assistant in processor design course

Georgia Institute of Technology *Jan. 2017 - May 2017*

- Helped students with processor design assignments on Altera FPGA boards.
- Graded exams and assignments

Teaching assistant in computer architecture course

Seoul National University *Sep. 2014 - Dec. 2014, Mar. 2015 - Jun. 2015*

- Developed lab assignments that implements a Y86 processor in Bluespec System Verilog.
- Verified the Y86 processor design on an FPGA.
- Helped developing course material.

Skills

Software languages	C/C++, JAVA, VBA, OCaml, Scheme, Matlab, and LaTeX
Machine Learning Frameworks	PyTorch
Formal verification tools	Coq Proof Assistant and ABC
Hardware languages	Verilog, Chisel, System C, and Bluespec System Verilog
ASIC synthesis tools	Synopsys Design Compiler and Cadence Encounter(Innovus)
FPGA synthesis tools	Xilinx Vivado and Altera Quartus
Parallel programming	OpenMP, OpenCL, and MPI
Image processing	OpenCV and Matlab
Languages	Korean, English, and Japanese