

# SEONJIN NA

*Phone:* +1 404-259-3240

*Email:* sna42@gatech.edu

*Github:* <https://github.com/seonjinna>

*Website:* <https://seonjinna.github.io>

## RESEARCH INTERESTS

---

I am a postdoctoral researcher at Georgia Tech supervised by Prof. Hyesoon Kim. Prior to joining Georgia Tech, I received a Ph.D. in Computer Science from KAIST (2023) advised by Prof. Jaehyuk Huh. My research interests lie in GPU architecture, trusted computing, heterogeneous systems, distributed computing, and systems for machine learning. During my Ph.D. studies, I focused on building secure architecture to provide a trusted execution environment(TEE) on accelerators such as GPUs, and NPUs with low-performance overhead. Currently, I am actively engaged in extending my research field to address various challenges in the Processing in Memory (PIM) and High-Performance Computing (HPC) systems.

## EMPLOYMENT

---

### Georgia Institute of Technology

*June. 2023 - present*

Postdoctoral Fellow, School of Computer Science

**Supervisor: Hyesoon Kim**

### Microsoft Research Asia

*Mar. 2019 - June. 2019*

Visiting Fellow

**Supervisors: Lintao Zhang & Yunxin Liu**

## EDUCATION

---

### KAIST

*Mar. 2018 - Feb. 2023*

Doctor of Philosophy, School of Computing

**Advisor: Jaehyuk Huh**

### KAIST

*Mar. 2016 - Feb. 2018*

Master of Science, School of Computing

**Advisor: Jaehyuk Huh**

### Sogang University

*Mar. 2012 - Feb. 2016*

Bachelor of Science, Computer Science

**Summa Cum Laude**

## RESEARCH PROJECTS

---

### Efficient On-chip Memory Management and Scheduling on NPUs

*Dec. 2021 - Present*

- Investigated the performance bottleneck of DNN Training on NPUs.
- Analyzed the data dependency of tensor computations in DNN training.
- Proposed an efficient on-chip memory management and scheduling policy to improve DNN training performance on NPUs.
- Participated as **second author** to conduct motivational experiments, discuss the idea, and help writing.
- **Published in MICRO 2023**

### Trusted Multi-GPU Architecture

*Sep. 2021 - Present*

- Investigated the performance overhead of prior secure communication techniques on a multi-GPU system.
- Analyzed the communication patterns of various GPU workloads and performance degradation of secure communication.
- Proposed an efficient data protection mechanism to minimize the performance overhead by secure communication
- Participated as **first author** to implement the simulator, conduct experiments, and lead the project.
- **Under Review**

#### **Efficient Memory Protection Mechanism for Secure NPU**

*Sep. 2020 - March. 2021*

- Investigated a significant performance degradation of CPU memory protection schemes on NPUs.
- Proposed a selective memory protection and multi-granular counter mode encryption techniques.
- Participated as **second author** to implement the simulator, discuss the idea, and conduct motivational experiments.
- Published in **ICCD 2022**

#### **Trusted NPU Architecture**

*Sep. 2019 - Sep. 2021*

- Designed a novel TEE design for secure NPU.
- Extended the existing CPU TEE design to isolate the NPU execution context from OS.
- Proposed a tree-less integrity protection by exploiting the characteristics of tensor-based NPU execution model.
- Participated as **third author** to implement the simulator, discuss the idea, and conduct motivational experiments.
- Published in **HPCA 2022**

#### **Efficient Memory Protection Mechanism for Secure GPU Memory**

*Sep. 2017 - Sep. 2020*

- Designed and implemented a secure GPU architecture that provides the confidentiality and integrity of the data on GPU memory with low-performance overhead.
- Analyzed the memory update behaviors of GPU benchmark suites and real-world GPU applications on Real-GPU hardware using NVbit tool.
- Proposed a efficient GPU memory protection technique to exploit the uniform memory update behavior of common GPU workloads.
- Participated as **first author** to implement the simulator, discuss the idea, conduct experiments, and lead the project.
- Published in **HPCA 2021**

#### **Machine Learning Inference on Mobiles**

*Mar. 2019 - Jun. 2019*

- Analyzed the performance characterization of mobile ML inferences using TensorFlow Lite framework.
- This project was done during Microsoft Research Asia internship.

#### **Hardware Prefetching**

*Mar. 2018 - Aug. 2018*

- Investigated and analyzed the performance of HW-based prefetching techniques on the CPU system.
- Implemented HW-based prefetching techniques on Gem5 simulator.

## **PUBLICATIONS**

- 
- Jungwoo Kim, **Seonjin Na**, Sanghyeon Lee, Sunho Lee, and Jaehyuk Huh, "Improving Data Reuse in NPU On-chip Memory with Interleaved Gradient Order for DNN Training", *the 56th IEEE/ACM International Symposium on Microarchitecture (MICRO)*, October 2023.
  - Sunho Lee, **Seonjin Na**, Jungwoo Kim, Jongse Park, and Jaehyuk Huh, "Tunable Memory Protection for Secure Neural Processing Units", *the 40th IEEE International Conference on Computer Design (ICCD)*, October 2022.

- Sunho Lee, Jungwoo Kim, **Seonjin Na**, Jongse Park, and Jaehyuk Huh, "TNPU: Supporting Trusted Execution with Tree-less Integrity Protection for Neural Processing Unit", *the 28th IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, February 2022.
- **Seonjin Na**, Sunho Lee, Yeonjae Kim, Jongse Park, and Jaehyuk Huh, "Common Counters: Compressed Encryption Counters for Secure GPU Memory", *the 27th IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, February 2021.

## RESEARCH EXPERIENCE

---

### Microsoft Research Asia

*Mar. 2019 - Jun. 2019*

- Research Intern, Advisor: Lintao Zhang, Yunxin Liu

### KAIST

*Mar. 2016 - Present*

- Graduate Research Assistant & Teaching Assistant

## PATENTS

---

### Dynamic One-time Pad Table Management for Secure Multi-GPU Communication

- Jaehyuk Huh, **Seonjin Na**, Jungwoo Kim, Sunho Lee
- Korea Patent; Pending

### Improving the Utilization of NPU On-chip Memory with Computation Rearrangement for DNN Training

- Jaehyuk Huh, Jungwoo Kim, **Seonjin Na**, Sanghyeon Lee, Sunho Lee
- Korea Patent; Pending

### Apparatus and Method for Providing Secure Execution Environment for NPU

- Jaehyuk Huh, Sunho Lee, **Seonjin Na**
- US Patent (With Samsung Electronics); Pending

### Hardware-based Security Architecture for Trusted Neural Processing Unit

- Jaehyuk Huh, Sunho Lee, **Seonjin Na**
- Korean Patent (With Samsung Electronics); Filling Date: 2021/07/23;

### Efficient Encryption Method and Apparatus for Hardware-based Secure GPU Memory

- Jaehyuk Huh, **Seonjin Na**, Sunho Lee, Yeonjae Kim, and Jongse Park
- Korea Patent; Filling Date: 2020/11/23; Issued Date: 2022/02/16

## AWARDS AND HONORS

---

### National Scholarship, KAIST

*Mar. 2016 - Present*

### Gold Prize

*Nov. 2015*

- The 2015 ACM-ICPC Asia Daejeon Regional Contest 4th place

### Honorable Mention

*Nov. 2013*

- The 2015 ACM-ICPC Asia Daejeon Regional Contest 13th place

### Academic Scholarship, 8 semesters

*Mar. 2012 - Sep. 2015*

- Sogang University

## ACADEMIC SERVICES

---

- **WebChair**: IEEE Computer Society TCuARCH
- **Reviewer**: IEEE Computer Architecture Letter 2023

## TEACHING EXPERIENCE

---

### KAIST

- CS230 System Programming: Fall 2016, Spring 2017, Fall 2018, Fall 2020
- CS311 Computer Organization: Fall 2019

### Sogang University

- Introduction to C Programming: Winter 2014

## SKILLS

---

- **Programming Languages** : C/C++, Go, CUDA, Python, Java
- **Library/Frameworks** : NVBit, Pytorch, Tensorflow
- **Simulators** : Fire-Sim, SST, GPGPU-Sim, MGPU-Sim, Gem5, Gem5-gpu, Scale-Sim