# Igjae Kim

Ph.D. Student, School of Computing, KAIST · Computer Architecture and Systems Lab (CASYS)
ijkim@kaist.ac.kr

I build system software that turns hardware characteristics into performance and resource efficiency. Working mainly in the Linux kernel and runtimes, I look for places where the assumptions of existing software no longer match the constraints of real hardware, and redesign memory management and execution environments around them. My current work centers on **CXL memory systems**, including memory pooling, sharing, and tiering on real CXL hardware, and on **infrastructure for LLM agents**, drawing on a background in hardware-based isolation (Intel SGX, confidential computing).

## EDUCATION

**Korea Advanced Institute of Science and Technology (KAIST)** *Sep. 2023 – Present*
Ph.D. in Computer Science, School of Computing · Advisor: Prof. Jaehyuk Huh *Daejeon, Korea*

**Korea Advanced Institute of Science and Technology (KAIST)** *Sep. 2021 – Aug. 2023*
M.S. in Computer Science, School of Computing · Advisor: Prof. Jaehyuk Huh *Daejeon, Korea*

**Ulsan National Institute of Science and Technology (UNIST)** *Mar. 2015 – Aug. 2021*
B.S. in Electrical and Electronic Engineering & Computer Science and Engineering (Double Major) *Ulsan, Korea*
Summa Cum Laude

## RESEARCH EXPERIENCE

### CXL MEMORY SYSTEMS

**Memory Pooling for Serverless Computing**
- Exploring how serverless runtimes can exploit memory that is shared across hosts through CXL.
- Building and evaluating system software for multi-host CXL memory on a real CXL prototype.

**Migration Control for Multi-Tenant Tiered Memory** *HPCA '27, arXiv '25*
- Observed that page migration in tiered memory does not always pay off: detecting and moving hot pages has its own cost, and some applications run faster with no migration at all.
- Designed and implemented a Linux kernel framework that (1) judges whether an application is migration-friendly from per-page *ping-pong* behavior, (2) detects access-pattern changes at low cost so that stopped migration can resume, and (3) controls migration per process in multi-tenant settings. Evaluated on a commercial CXL-based tiered memory system (arXiv '25).
- Co-authored NEST, multi-tenant tiered memory management with near-memory demand estimation (HPCA '27).

**Huge Page Management for Disaggregated Memory**
- Huge pages cut address-translation cost but waste capacity-limited local memory through internal fragmentation. Designed a page-migration scheme in the Linux memory-management path that balances the two.
- Reduced TLB misses by 75% on average compared with 4 KB base pages. With only 30% of memory local, outperformed an all-local 4 KB configuration by 2.3%.

**ShieldCXL: Access-Pattern Protection with Sealed CXL Memory** *TACO '25*
- Protects the DDR interface of a CXL memory board with a tamper-responding enclosure, removing the need for DDR-level integrity trees and ORAM.
- Secures the CXL link instead: encrypted flit counters and MACs block injection and replay, and dummy slots hide request types and access timing.
- 9.16× faster than DDR-based Path ORAM, with only a 4.7% slowdown over unprotected CXL memory in simulation.

### AI SYSTEMS & SECURITY

**Sandbox Infrastructure for LLM Agents on CXL Memory Pools**
- Building scalable, resource-efficient sandbox infrastructure for LLM agents that execute code and tools, on memory pooled across hosts through CXL.
- Extends my CXL memory pooling work from serverless functions to agent workloads.

**Interference-Aware DNN Serving on Heterogeneous Edge Processors** *ICCD '24*
- Characterized interference among DNNs running concurrently on the GPU and two DLAs of an NVIDIA Jetson AGX Xavier, caused by shared memory bandwidth and by DLA layers falling back to the GPU.
- A lightweight MLP interference predictor (95.9% accuracy) drives a scheduler that improves goodput by 40% and SLO-bounded throughput by 36.1% on average over an affinity-based baseline.
- My role: interference profiling and experiment design. Presented the paper at ICCD 2024.

**TWEEZER: A Secure Persistent Key-Value Store on Intel SGX** *FAST '22*

- First author; led the research and paper writing.
- Prior SGX key-value stores keep MAC metadata that grows with the data in the EPC, causing heavy paging.
- Exploited LSM-tree properties: per-SSTable MAC keys that remove the Merkle tree, per-key-value MACs with key-ordering checks that move the block cache out of the EPC, and hash-chained logs that replace slow SGX trusted counters. Built on RocksDB.
- 1.94–6.23× higher throughput than Speicher with 5.24–7.57× less EPC paging.

**SecPIMlet: Confidential Computing for Shared Processing-in-Memory** *MICRO '26*

- Co-authored work that protects data and execution state in spatio-temporally shared processing-in-memory.

## PUBLICATIONS

[1] **NEST: Multi-Tenant Tiered Memory Management with Near Memory Demand Estimation**
Kwanghoon Choi, **Igjae Kim**, Hongjin Kim, Suhwan Kim, and Jaehyuk Huh
*IEEE International Symposium on High-Performance Computer Architecture (HPCA), 2027*

[2] **SecPIMlet: Confidential Computing for Spatio-temporally Shared Processing-in-Memory**
Sunho Lee, Minwoo Noh, Jeongwon Choi, Kwanghoon Choi, **Igjae Kim**, Chang Hyun Park, and Jaehyuk Huh
*IEEE/ACM International Symposium on Microarchitecture (MICRO), 2026*

[3] **Adaptive Migration Decision for Multi-Tenant Memory Systems**
Hyungjun Cho, **Igjae Kim**, Kwanghoon Choi, Hongjin Kim, Wonjae Lee, Junhyeok Im, Jinin So, and Jaehyuk Huh
*arXiv preprint arXiv:2505.09164, 2025*

[4] **ShieldCXL: A Practical Obliviousness Support with Sealed CXL Memory**
Kwanghoon Choi, **Igjae Kim**, Sunho Lee, and Jaehyuk Huh
*ACM Transactions on Architecture and Code Optimization (TACO), Vol. 22, No. 1, 2025*

[5] **Interference-Aware DNN Serving on Heterogeneous Processors in Edge Systems**
Yeonjae Kim, **Igjae Kim**, Kwanghoon Choi, Jeongseob Ahn, Jongse Park, and Jaehyuk Huh
*IEEE International Conference on Computer Design (ICCD), pp. 199–206, 2024* (presenter)

[6] **A Log-Structured Merge Tree-aware Message Authentication Scheme for Persistent Key-Value Stores**
**Igjae Kim**, J. Hyun Kim, Minu Chung, Hyungon Moon, and Sam H. Noh
*USENIX Conference on File and Storage Technologies (FAST), pp. 363–380, 2022* (first author, presenter)

## EXPERIENCE

**CryptoLab Inc.** *Jul. 2021 – Aug. 2021*
Research Intern

- Implemented CKKS homomorphic encryption and decryption inside an Intel SGX enclave for secure DNN inference.

**Ulsan National Institute of Science and Technology (UNIST)** *Spring 2020*
Undergraduate Research Intern · Advisor: Prof. Hyungon Moon · Led to TWEEZER (FAST '22)

## AWARDS & HONORS

- **LS Award**, UNIST *Aug. 2021*
- **Summa Cum Laude**, UNIST *Aug. 2021*
- **Dean's List**, UNIST (Spring & Fall 2015, Spring 2016, Fall 2019, Spring & Fall 2020) *2015 – 2020*

## TEACHING ASSISTANT

- CS230 / CS20300 System Programming, KAIST *Fall 2021, 2022, 2023, 2025*
- CS30101 Computer Organization, KAIST *Spring 2025*
- CS492 Special Topics in Computer Science: Parallel Computing, KAIST *Fall 2024*

## SKILLS

**Programming** C/C++, Python, Scala
**Systems** Linux kernel, Apache OpenWhisk, RocksDB, NVIDIA Jetson, Intel SGX SDK