Byungsoo Oh

byungsoo@cs.cornell.edu | byungsoo-oh.github.io

RESEARCH INTERESTS

Systems for ML, Hardware-Software Co-Design, Networking for AI Infrastructure

EDUCATION

Cornell University — Ithaca, NY, USA
Ph.D. in Computer Science — Aug 2024 – Present
Advisor: Prof. Rachee Singh

Korea Advanced Institute of Science and Technology (KAIST) — Daejeon, South Korea
M.S. in Computer Science — Mar 2018 – Feb 2020

Sogang University — Seoul, South Korea
B.S. in Computer Science and Engineering — Mar 2012 – Feb 2018
Graduated with honors, *Summa Cum Laude*

PROFESSIONAL EXPERIENCE

Microsoft Research — Redmond, WA, USA
Research Intern — May 2025 – Aug 2025
Mentors: Dr. Aashaka Shah, Dr. Abhinav Jangda, Dr. Roshan Dathathri, and Prof. Alvin Lebeck

Samsung Research — Seoul, South Korea
Research Engineer — Feb 2020 – Jun 2024

PUBLICATIONS

- Osayamen Jonathan Aimuyo, **Byungsoo Oh**, Rachee Singh, "FlashMoE: Fast Distributed MoE in a Single Kernel", Conference on Neural Information Processing Systems (**NeurIPS**), San Diego, CA, USA, 2025
- Taegeon Um*, **Byungsoo Oh***, Minyoung Kang*, Woo-Yeon Lee, Goeun Kim, Dongseob Kim, Youngtaek Kim, Mohd Muzzammil, Myeongjae Jeon, "Metis: Fast Automatic Distributed Training on Heterogeneous GPUs", USENIX Annual Technical Conference (**USENIX ATC**), Santa Clara, CA, USA, 2024 (* *denotes co-first authors*)
- Taegeon Um, **Byungsoo Oh**, Byeongchan Seo, Minhyeok Kweun, Goeun Kim, Woo-Yeon Lee, "FastFlow: Accelerating Deep Learning Model Training with Smart Offloading of Input Data Pipeline", International Conference on Very Large Data Bases (**VLDB**), Vancouver, Canada, 2023
- Minhyeok Kweun, Goeun Kim, **Byungsoo Oh**, Seongho Jung, Taegeon Um, Woo-Yeon Lee, "PokéMem: Taming Wild Memory Consumers in Apache Spark", IEEE International Parallel and Distributed Processing Symposium (**IPDPS**), Lyon, France, 2022
- Seungju Cho, Tae Joon Jun, **Byungsoo Oh**, Daeyoung Kim, "DAPAS: Denoising Autoencoder to Prevent Adversarial attack in Semantic Segmentation", International Joint Conference on Neural Networks (**IJCNN**), Glasgow, UK, 2020
- **Byungsoo Oh**, Daeyoung Kim, "Serverless-Enabled Permissioned Blockchain for Elastic Transaction Processing", ACM/IFIP International Middleware Conference (**Middleware**), *Poster Paper*, Davis, CA, USA, 2019
- **Byungsoo Oh**, Tae Joon Jun, Wondeuk Yoon, Yunho Lee, Sangtae Kim, Daeyoung Kim, "Enhancing Trust of Supply Chain Using Blockchain Platform with Robust Data Model and Verification Mechanisms", IEEE International Conference on Systems, Man, and Cybernetics (**SMC**), Bari, Italy, 2019

PATENTS

- Minyoung Kang, **Byungsoo Oh**, Taegeon Um, "Method and System for Elastic Knowledge Distillation with Adaptive Coordination", US Patent, US20250068943A1, Published: Feb 27, 2025
- Taegeon Um, Minhyeok Kweun, **Byungsoo Oh**, "Smart Offloading for AI Input Data Pipeline Acceleration", US Patent, US20240135189A1, Published: Apr 25, 2024

- Minyoung Kang, **Byungsoo Oh**, Taegeon Um, "Device Placement Strategies for Optimizing 3D Parallelism in Non-Uniform Topology Environments", US Patent, Pending, 2023
- Daeyoung Kim, **Byungsoo Oh**, "Method and System for Enhancing Trust of Supply Chain Using Blockchain Platform with Robust Data Model and Verification Mechanisms", Korean Patent, No. 10-2620822-0000, Issued: Dec 2023

HONORS AND AWARDS

- **LinkedIn Fellowship** 2025–2026
  Cornell Bowers CIS-LinkedIn grant for academic year of 2025-2026
- **USENIX ATC 2024 Student Grant** 2024
  Travel grant awarded to attend USENIX ATC 2024 (co-located with OSDI 2024) in Santa Clara
- **National Full Scholarship**, Korea Ministry of Science and ICT 2018–2020
- **Award for Top 1% Students in College of Engineering (Dean's List)**, Sogang University 2017
  2 semesters (Spring 2017, Fall 2017)
- **Academic Excellence Scholarship**, Sogang University 2013–2017
  6 semesters (Spring 2013, Fall 2015, Spring 2016, Fall 2016, Spring 2017, Fall 2017)

ACADEMIC SERVICE

- **Artifact Evaluation Committee:** SOSP 2025

TEACHING EXPERIENCE

- TA, Introduction to Computer Networks (CS4450/5456), Cornell University Fall 2024
- TA, Introduction to System Programming (CS230), KAIST Spring 2019, Spring 2018
- TA, Embedded Operating Systems (CS632), KAIST Fall 2018

TECHNICAL SKILLS

- **Programming Languages:** C, C++, Python, Go, JavaScript, Java, Scala, Markdown, LaTeX
- **ML Systems:** PyTorch, TensorFlow, DeepSpeed, Megatron-LM, Alpa, vLLM, verl, CUDA, NSight Systems & Compute
- **Infra & Tools:** Docker, Kubernetes, gRPC, Apache Airflow, Apache Spark, Apache Hadoop

OPEN SOURCE CONTRIBUTIONS

- **DeepSpeed.** Bug Fix [issue] [code]
- **TensorFlow.** Documentation improvement for `tf.data service` [code]
- **Apache Spark.** Bug Fix [code], Benchmark [code]

LANGUAGES

Korean (*native*), English (*fluent*)