

# Jae-Won Chung

☎ +1 (734) 496-1803 | ✉ jwncung@umich.edu | 🌐 jaewonchung.me | 📺 jaywonchung | 📄 jae-won-chung-cs

## Summary

I am a fifth year PhD candidate in CSE at the University of Michigan, advised by [Professor Mosharaf Chowdhury](#). I build **efficient software systems for deep learning**, with a recent focus on the efficient management of not only time, but also **energy**. My research views energy as a first-class systems resource that is worth carefully optimizing and allocating based on precise measurement and understanding.

I am passionate about **open-source** software and making real-world impact with my research. My open-source works, including the [Zeus](#) library, have received wide recognition from academia and industry from for instance [Google](#), [PyTorch Foundation](#), and [GitHub](#). I created and lead the [ML.ENERGY initiative](#) as part of my research and open-source efforts, which is now a cross-institutional effort.

## Education

### University of Michigan

PH.D. CANDIDATE IN COMPUTER SCIENCE AND ENGINEERING

*Ann Arbor, MI, USA*

*Sep 2021 - present*

### University of Michigan

M.S. IN COMPUTER SCIENCE AND ENGINEERING

*Ann Arbor, MI, USA*

*Sep 2021 - Apr 2023*

### Seoul National University

B.S. IN ELECTRICAL AND COMPUTER ENGINEERING

*Seoul, South Korea*

*Mar 2015 - Aug 2021*

- GPA: 4.04/4.3 (overall) 4.15/4.3 (major), Summa Cum Laude. Period includes two years of military service.

## Publications

### Peer-reviewed conference publications

\* Equal contribution

- **The ML.ENERGY Benchmark: Toward Automated Inference Energy Measurement and Optimization**, Jae-Won Chung, Jeff J. Ma, Ruofan Wu, Jiachen Liu, Oh Jun Kweon, Yuxuan Xia, Zhiyu Wu, Mosharaf Chowdhury, **NeurIPS Datasets & Benchmarks track (spotlight)**, 2025 (Spotlight acceptance rate = 2.81%)
- **Reducing Energy Bloat in Large Model Training**, Jae-Won Chung, Yile Gu, Insu Jang, Luoxi Meng, Nikhil Bansal, Mosharaf Chowdhury, **SOSP**, 2024 (Acceptance rate = 17.34%)
- **Zeus: Understanding and Optimizing GPU Energy Consumption of DNN Training**, Jie You\*, Jae-Won Chung\*, Mosharaf Chowdhury, **NSDI**, 2023 (Acceptance rate = 18.38%)
- **ShadowTutor: Distributed Partial Distillation for Mobile Video DNN Inference**, Jae-Won Chung, Jae-Yun Kim, Soo-Mook Moon, International Conference on Parallel Processing (**ICPP**), 2020 (Acceptance rate = 28.99%)

### Preprints and workshop publications

- **GPU-to-Grid: Voltage Regulation via GPU Utilization Control**, Zhirui Liang, Jae-Won Chung, Mosharaf Chowdhury, Jiasi Chen, Vladimir Dvorkin, 2026
- **Where Do the Joules Go? Diagnosing Inference Energy Consumption**, Jae-Won Chung, Ruofan Wu, Jeff J. Ma, Mosharaf Chowdhury, 2026
- **Kareus: Joint Reduction of Dynamic and Static Energy in Large Model Training**, Ruofan Wu, Jae-Won Chung, Mosharaf Chowdhury, 2026
- **Cornserve: Efficiently Serving Any-to-Any Multimodal Models**, Jeff J. Ma\*, Jae-Won Chung\*, Akshay Jajoo, Myungjin Lee, Mosharaf Chowdhury, 2025
- **Toward Cross-Layer Energy Optimizations in AI Systems**, Jae-Won Chung, Nishil Talati, Mosharaf Chowdhury, **DOE ASCR Energy-Efficient Computing for Science Workshop**, 2024
- **Andes: Defining and Enhancing Quality-of-Experience in LLM-Based Text Streaming Services**, Jiachen Liu, Jae-Won Chung, Zhiyu Wu, Fan Lai, Myungjin Lee, Mosharaf Chowdhury, 2024
- **Chasing Low-Carbon Electricity for Practical and Sustainable DNN Training**, Zhenning Yang, Luoxi Meng, Jae-Won Chung, Mosharaf Chowdhury, **ICLR Workshop: Tackling Climate Change with Machine Learning**, 2023

## Open Source Projects

- **Zeus** (☆336 ♪ 40), A framework for deep learning energy measurement and optimization. PyTorch ecosystem project. Provides a unified abstraction over various CPU, GPU, and SoC vendors, and an array of time & energy optimizers.
- **Cornserve** (☆111 ♪ 6), An efficient serving platform for any-to-any multimodal models. Takes a microservice approach to serving complex-structured multimodal input & output models.
- **OpenG2G** (☆2 ♪ 0), Modular simulation library for AI datacenter-grid interaction. Enables research on how AI datacenters interact with the power grid through composable, extensible simulation modules.
- **BERT4Rec-VAE-Pytorch** (☆408 ♪ 98), A PyTorch framework for recommendation model training, with abstract classes for pluggable model, dataset, and samplers. BERT4Rec and Netflix VAE models implemented.
- **Reason** (☆197 ♪ 4), A shell for managing research papers, written in Rust. Supports importing papers from file and URL, attaching markdown notes, and creating an HTML book with notes. Uses serde to persist data in human-readable and cloud sync-friendly format.
- **Pegasus** (☆32 ♪ 3), An SSH command runner with a focus on simplicity, written in Rust. Runs multiple commands asynchronously using the tokio runtime and streams stdout and stderr back to the user. Battle-tested through multiple research projects and benchmarking.

Number of stars and forks are up-to-date as of February 25th, 2026.

## Honors & Awards

Nov 2025	<b>Slingshot Grant</b> , \$25,000 USD AWS Credits for the development of the <u>ML.ENERGY Initiative</u>	<i>Laude Institute</i>
Aug 2025	<b>GitHub Secure Open Source Fund</b> , \$10,000 for the development of the <u>Zeus</u> project	<i>GitHub</i>
May 2024	<b>PyTorch Ecosystem Project</b> , <u>Zeus</u> was included in the PyTorch Ecosystem	<i>PyTorch Foundation</i>
Jan 2024	<b>Research award</b> , \$20,000 for the development of the <u>ML.ENERGY Initiative</u>	<i>Salesforce</i>
Jan 2024	<b>Mozilla Technology Fund 2024</b> , \$50,000 for the development of the <u>Zeus</u> project	<i>Mozilla</i>
Nov 2022	<b>Carbon Hack '22 Second Best Solution</b> , <u>Carbon-Aware DNN Training with Zeus</u> , \$25,000	<i>Green Software Foundation</i>
Jul 2021	<b>Kwanjeong Overseas Scholarship</b> , \$25,000	<i>Kwanjeong Educational Foundation</i>
Mar 2019	<b>Kwanjeong Undergraduate Scholarship</b> , \$20,000 over two years	<i>Kwanjeong Educational Foundation</i>

## Talks

Dec 2025	<b>Energy and Power as First-Class ML Design Metrics (w/ NVIDIA, Google, Meta)</b>	<i>NeurIPS 25 Tutorial</i>
Oct 2025	<b>Energy as a First-Class ML Design Metric</b>	<i>UW-Madison MadSystems Seminar</i>
Sep 2025	<b>Power and Energy as First-Class AI Design Metrics</b>	<i>KPAI (Bay Area Korean AI Meetup)</i>
Jun 2025	<b>Energy as a First-Class Resource in Machine Learning Systems</b>	<i>Pruna AI</i>
May 2025	<b>Energy-Efficient Systems for Machine Learning</b>	<i>Harvard Power and AI Initiative</i>
Nov 2024	<b>Energy-Efficient Systems for Machine Learning</b>	<i>SOSP 24 Doctoral Workshop</i>
Apr 2024	<b>Power and Energy Considerations in Machine Learning Systems</b>	<i>University of Michigan (EECS 598)</i>
Oct 2023	<b>Energy-Efficient Software Systems for Machine Learning</b>	<i>Seoul National University</i>
Oct 2023	<b>Energy-Efficient Deep Learning with PyTorch and Zeus</b>	<i>PyTorch Conference</i>
Sep 2023	<b>Energy-Efficient Deep Learning with Zeus</b>	<i>Massachusetts Institute of Technology</i>

## Selected Media Coverage

My research and open-source works have been covered by various media outlets, including MIT Technology Review, ArsTechnica, and Science News.

Jul 2025	<b><u>How much energy does your AI prompt use? It depends.</u></b>	<i>Science News</i>
May 2025	<b><u>We did the math on AI's energy footprint. Here's the story you haven't heard.</u></b>	<i>MIT Technology Review</i>
May 2025	<b><u>AI Consumes Lots of Energy. Can It Ever Be Sustainable?</u></b>	<i>The New Stack</i>
Mar 2025	<b><u>Can we make AI less power-hungry? These researchers are working on it.</u></b>	<i>ArsTechnica (front page)</i>
Nov 2024	<b><u>Up to 30% of the power used to train AI is wasted: A software tool could help fix that.</u></b>	<i>Tech Xplore</i>
Apr 2023	<b><u>University of Michigan's 'Zeus' Framework Downsizes AI's Massive Carbon Footprint.</u></b>	<i>HPCWire</i>
Apr 2023	<b><u>Researchers claim they can cut AI training energy demands by 75%.</u></b>	<i>DatacenterDynamics</i>

## Service

---

- **Ad-hoc reviewer**, Nature
- **Systems Reading Group Organizer**, Paper reading group inside Michigan CSE. Organizer from September 2022 to December 2025.

## Teaching

---

- **Systems for Generative AI (UMich, Fall 25)**, GSI, three lectures on GenAI and GenAI systems fundamentals. **Instructor evaluation median 5.00 (highest possible)**. Written comments included phrases “excellent” (three times), “perfect,” and “one of the best.”
- **Operating Systems (SNU, Spring 21)**, Lead TA, managed Linux kernel hacking projects and led student team design reviews.
- **Computer Architecture (SNU, Fall 20)**, Peer tutor, provided 30 hours of online lecture. **Best Tutor Award**.