

# Daniar H. Kurniawan

daniar@uchicago.edu • [LinkedIn](#) • [GitHub](#) • [Google Scholar](#) • [Homepage](#)

## EDUCATION

---

### UNIVERSITY OF CHICAGO, USA

Ph.D. Cand. in Computer Science (Area: **AI/ML and Storage Systems Architecture**) June 2024

- CERES outstanding 1<sup>st</sup> year student award recipient
- University Unrestricted (UU) Fellowship recipient

M.Sc. in Computer Science (Area: Storage Systems Runtime) December 2020

### BANDUNG INSTITUTE OF TECHNOLOGY (ITB), Indonesia

July 2017

B. Eng. in Computer Science (Area: Algorithm and Software Engineering)

- Dean's List Academic Achievement Award recipient
- Merit-based full academic scholarship recipient

## WORK EXPERIENCE

---

### UNIVERSITY OF CHICAGO – Chicago, IL, USA

August 2018 – June 2024

Graduate Research Assistant, UCARE Group

- Built a novel caching algorithm into PyTorch reducing Meta's Deep Learning memory usage by up to 94%
- Optimized I/O admission control algorithm to reduce p99 tail latency by up to 90%
- Introduced a novel inference method with a 10x latency speedup compared to the state-of-the-art
- Improved Neural Network deployment in Linux Kernel, achieving a 10-microsecond inference latency
- Designed and built an ML-based drift detection algorithm that outperforms statistical algorithms
- Developed and integrated an ML model into Ceph distributed storage and evaluated it on 20 nodes cluster

### SEAGATE TECHNOLOGY – Fremont, CA, USA

Summer of 2021, 2022, and 2023

Intern, Seagate Research Group

- Developed ML-based prefetcher for hard-disk firmware that achieved an 80% hit rate on real-world traces
- Performed scalability evaluation of CORTX, Ceph, MinIO, and Cassandra on a cluster of 8 nodes
- Fixed 10+ issues within CORTX modules and submitted 10+ pull request to its GitHub repository

### VMWARE – Palo Alto, CA, USA

June 2020 – September 2020

Intern, VMware Research Group

- Studied ~300 of Cassandra's and Hive's source code files to understand the data path
- Integrated ~10,000 of lines of Cassandra's and Hive's data processing codes into Hillview
- Modified Hillview web-based UI platform to setup cluster configuration
- Improved big data analysis speed by up to 10x; tested it on a cluster of 20 nodes

### MICROSOFT – New York City, NY, USA

November 2018

Visiting student, Microsoft Research

- Analyzed ~1TB worth of data collected from Microsoft cloud systems
- Co-authored E2E paper that introduces a better resource allocation that increases user engagement by 28%
- Prototyped the E2E integration into Cassandra and RabbitMQ cluster

## AWARDS

---

**Travel Grants:** ICEEP'15, ICAICTA'16, ICEEP'17, SOCC'18, SYSTOR'22 2015 – 2022

**Third place award at CERN annual Webfest, Switzerland** 2017

**First place award at IBM business plan competition, Indonesia** 2015

## TECHNICAL SKILLS

---

**Programming Language** : Python, C/C++, Java, JavaScript  
**ML (Machine Learning)** : Keras, PyTorch, TensorFlow  
**Testbed/Cloud Platform** : Amazon Web Service, Chameleon, Emulab, Google Cloud Platform  
**Complex Systems Hacking** : Cassandra, Couchbase, Hadoop, Hive, JVM, Linux Kernel, PyTorch, Spark  
**Database Systems** : Advanced user of Cassandra, Hadoop, Hive, Ceph, MinIO, MongoDB, Spark

## SOFTWARE (Open-source contribution)

---

**EVSTORE**: <https://github.com/ucare-uchicago/ev-store-dlrm> (24☆) 2023  
**Developed and integrated an optimized caching layer into PyTorch** embedding lookup module.  
Written in ~9,000 LOC (70% Python, 25% C/C++, and 5% bash scripts)

**CHAMELEON TROVI**: <https://chameleoncloud.org/experiment/share> 2022 – 2023  
**Published 7 highly reproducible artifacts** that have been launched 90+ times by international researchers

**CORTX**: <https://github.com/seagate/cortx> (632☆) 2021 – 2022  
**Fixed 10+ issues, submitted 10+ pull requests**, and improved the clarity of various readme files

**HILLVIEW**: <https://github.com/vmware-archive/hillview> (99☆) 2020  
**Added ~10,000 of lines of code to enable parallel data processing** when connected to Cassandra and Hive clusters. Enabled over 10x faster big data analysis on both clusters.

**DMCK**: <https://github.com/ucare-uchicago/dmck> (5☆) 2017 – 2018  
**Developed a module to detect distributed-concurrency bugs** in Spark and improved the multi-threading implementation that speeds up the node restart by 4x

## INVITED TALKS

---

**SEAGATE TECHNOLOGY** – Fremont, CA, USA June 2023  
The Design and Development of ML-based Prefetcher

**SEAGATE TECHNOLOGY** – Fremont, CA, USA April 2023  
EVStore: Storage and Caching Capabilities for Scaling EV Tables in Deep Recommendation Systems.

**META** – Menlo Park, CA, USA April 2023  
EVStore: Storage and Caching Capabilities for Scaling EV Tables in Deep Recommendation Systems.

## SELECTED PUBLICATION ([Google Scholar](#): 12 Papers)

---

1. [D. H. Kurniawan](#), R. Wang, K. Zulkifli, F. Wiranata, J. Bent, Y. Vigfusson, H. S. Gunawi. “**EVStore: Storage and Caching Capabilities for Scaling Embedding Tables in Deep Recommendation Systems.**” *In Proceedings of the 28<sup>th</sup> ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ACM ASPLOS)*, 2023.
2. M. Wang, C. Stuardo, [D. H. Kurniawan](#), R. A. O. Sinurat, H. S. Gunawi. “**Layered Contention Mitigation for Cloud Storage.**” *In Proceedings of the IEEE International Conference on Cloud Computing (IEEE CLOUD)*, 2022.
3. X. Zhang, S. Sen, [D. H. Kurniawan](#), H. S. Gunawi, and Junchen Jiang. “**E2E: Embracing User Heterogeneity to Improve Quality of Experience on the Web.**” *In Proceedings of the ACM Special Interest Group on Data Communication (ACM SIGCOMM)*, 2019.
4. J. F. Lukman, H. Ke, C. A. Stuardo, R. O. Suminto, [D. H. Kurniawan](#), D. Simon, S. Priambada, C. Tian, F. Ye, T. Leesatapornwongsa, A. Gupta, S. Lu, H. S. Gunawi. “**FlyMC: Highly Scalable Testing of Complex Interleavings in Distributed Systems.**” *In Proceedings of The European Conference on Computer Systems (ACM EUROSYS)*, 2019.