

# Yuheng Zhao

yzhaoep@connect.ust.hk

+86 199 2126 2050

HKUST, Clear Water Bay, Kowloon, Hong Kong

<https://github.com/Farrrrland>

## ABOUT ME

---

I'm a 2<sup>nd</sup> year Ph.D. student at HKUST, supervised by Prof. Wei WANG. My research interest is Cloud Computing and Serverless Machine Learning Systems.

## EDUCATION

---

**Fudan University** | CS | BSc

Sep 2018 - Jun 2022

**Coursework:** Data Structure, Computer Architecture, Database, Operating Systems, Probability & Mathematical Statistics, Computer Networks, Distributed Systems, Algorithms, Big Data Systems.

- Research assistant at Cryptography and Information Security Laboratory, supervised by Prof. Haibin Kan.

**The Hong Kong University of Science and Technology** | CSE | Ph.D

Sep 2022 - Now

## INTERNSHIP

---

**Microsoft Research Asia** - Research Intern - Beijing

Nov 2021 - Jun 2022

- Worked as research intern in Data, Knowledge and Intelligence group at MSRA.

- Participated in a workload optimization project on Azure Cloud, including Region Shift and Workload Profiling to improve the compute resource utilization and overall health condition in different regions.

- Analysis and characterizing data on Azure, comparing different patterns of private/public cloud and published the work on DSN'23.

- Awarded "The Star of Tomorrow".

**Ant Group** - Software Engineer Intern - Shanghai

Jul 2021 - Oct 2021

- Supported the data service system "Turing" for Alipay, one of the largest mobile payment platforms in the world.

- Developed the asset management part of "Turing", improved the process of upgrading labels and developed a label comparison process to enable users to access data easier and faster.

**University of Chicago, Globus Lab** - Research Assistant - Chicago

May 2021 - Oct 2021

- Participated in a research project on "FuncX", a federated FaaS platform, focused on the container management algorithm of FuncX container service.

- Made analysis and optimization of "LANDLORD", the currently used algorithm, improved the performance of the container service.

## SELECTED ACADEMIC PROJECTS

---

**Making RL Fast as Possible, Serverless when Necessary**

Apr 2023 - Now

- We developed a RL system adopting serverless features.

- Our system achieve ~50% resource saving with comparable performance (system throughput difference within 10%) compared with SOTA solution, Ray.

- Our work is targeting NSDI'25.

**Data Object Management Strategy on Serverless Platform**

Mar 2022 - May 2022

- Focused on data object management strategy on Pheromone, a data-centric orchestration serverless platform.

- Implemented a data object life-cycle management and garbage collection strategy with high efficiency and flexibility.

## PUBLICATIONS

---

X. Qin, M. Ma, Y. Zhao, J. Zhang, et al. How Different are the Cloud Workloads? Characterizing Large-Scale Private and Public Cloud Workloads. DSN'23, Research Track.