Zhinan Cheng

Senior Software Engineer, Tech Lead | TikTok | London, UK

chingchinam@gmail.com | zncheng.github.io

Google Scholar | GitHub

Research interests

Reliable distributed systems, cloud infrastructure, automated healing, and fault-tolerant stream processing. Other research includes disaggregated memory and energy-efficient mobile computing.

Professional experience

TikTok | London, UK — December 2024–present

Senior Software Engineer, Tech Lead

- Lead platform management and reliability development within the multi-cloud CDN team responsible for **all TikTok traffic worldwide**.
- Develop AI-assisted observability, root-cause analysis, and self-healing for tens of thousands of overseas domains.
- Standardize domain lifecycle management and multi-provider security orchestration; turn alert configuration and disaster-recovery failover into self-service workflows.

Tencent Cloud | Shenzhen, China — April 2022–December 2024

Senior Software Engineer

- Led the design of Klever, a service mesh deployed across **more than 1,500 edge data centers worldwide**, supporting service discovery, traffic management, and tracing for Tencent Serverless@Edge.
- Designed CDN cache self-healing that automates detection, root-cause diagnosis, and repair of failed and slow-performing cache nodes.

Alibaba Cloud | Hangzhou, China — November 2020–January 2022

Research Intern

- Co-developed AIHS, deployed across **approximately 600,000 Alibaba Cloud servers**. In a seven-month production evaluation, AIHS repaired 92.4% of 33.7 million failures and reduced average downtime per failed server by 51% versus policy-based self-healing. Published at SRDS 2021 as co-first author.
- Developed DRAM-error-based failure prediction using an eight-month dataset from 250,000 servers with over 3 million memory modules. Published at SRDS 2022 with public code and data.

Education

The Chinese University of Hong Kong — 2022

Ph.D., Computer Science and Engineering. Advisor: Patrick P. C. Lee.

University of Science and Technology of China — 2017

M.Eng., Computer Science and Technology. Advisor: Xi Li.

South China Normal University — 2014

B.Eng., Computer Science and Technology.

Publications

An asterisk denotes equal contribution.

Distributed systems and cloud infrastructure

1. Pulse: Fine-Grained Hierarchical Hashing Index for Disaggregated Memory.
   Guangyang Deng, Zixiang Yu, Zhirong Shen, Qiangsheng Su, **Zhinan Cheng**, and Jiwu Shu.
   IEEE HPCA, 2026. PDF · Code
   ICORE 2026: A* · CCF: A · Acceptance rate: 21.1% (119/563)

2. An In-Depth Correlative Study Between DRAM Errors and Server Failures in Production Data Centers.
   Zhinan Cheng, Shujie Han, Patrick P. C. Lee, Xin Li, Jiongzhou Liu, and Zhan Li.
   IEEE SRDS, 2022. PDF · Code · Dataset
   Acceptance rate: 22.9% (24/105)

3. Enabling Low-Redundancy Proactive Fault Tolerance for Stream Machine Learning via Erasure Coding.
   Zhinan Cheng, Lu Tang, Qun Huang, and Patrick P. C. Lee.
   IEEE SRDS, 2021. PDF · Code
   Research-paper acceptance rate: 23.5% (as reported)

4. Automated Intelligent Healing in Cloud-Scale Data Centers.
   Rui Li*, **Zhinan Cheng***, Patrick P. C. Lee, Pinghui Wang, Yi Qiang, Lin Lan, Cheng He, Jinlong Lu, Mian Wang, Xinquan Ding.
   IEEE SRDS, 2021. PDF · Code
   Practical Experience Report

5. On the Performance and Convergence of Distributed Stream Processing via Approximate Fault Tolerance.
   Zhinan Cheng, Qun Huang, Patrick P. C. Lee.
   The VLDB Journal, 28(5), 2019. DOI · PDF · Code
   CCF: A · Journal Impact Factor: 5.3 (2025)

Energy efficient mobile computing

6. Building A Game Benchmark for Cooperative CPU-GPU with Pseudo User-interaction.
   Zhen Wang*, **Zhinan Cheng***, Xi Li, Chao Wang, Xianglan Chen and Xuehai Zhou.
   IEEE ISPA, 2017. PDF

7. Rethinking Energy-Efficiency of Heterogeneous Computing for CNN-Based Mobile Applications.
   Zhen Wang, Xi Li, Chao Wang, **Zhinan Cheng**, Jiachen Song and Xuehai Zhou.
   IEEE ISPA, 2017. Short paper. PDF

8. Behavior-Aware Integrated CPU-GPU Power Management For Mobile Games.
   Zhinan Cheng, Xi Li, Beilei Sun, Jiachen Song, Chao Wang and Xuehai Zhou.
   IEEE MASCOTS, 2016. Short paper. PDF

Publications continued

9. Display Power Reduction for Mobile Close-Source Game.
   Zhinan Cheng, Xi Li, Jiachen Song, Beilei Sun, Xuehai Zhou and Chao Wang.
   IEEE ASAP, 2016. Poster. PDF

10. Improving Energy Efficiency by Phase-Grained Migration for Asymmetric Multicore.
    Youjun Xu, Xianglan Chen, **Zhinan Cheng**, et al.
    IEEE ISPA, 2016. PDF

11. FCM: Towards Fine-Grained GPU Power Management for Closed Source Mobile Games.
    Jiachen Song, Xi Li, Beilei Sun, **Zhinan Cheng**, Chao Wang and Xuehai Zhou.
    ACM/IEEE GLSVLSI, 2016. Poster. PDF

12. Automatic Frame Rate-based DVFS of Game.
    Zhinan Cheng, Xi Li, Beilei Sun, Ce Gao, Jiachen Song.
    IEEE ASAP, 2015. Poster. PDF

13. Texture-Directed Mobile GPU Power Management for Closed-Source Games.
    Beilei Sun, Xi Li, Jiachen Song, **Zhinan Cheng**, Yuan Xu and Xuehai Zhou.
    IEEE HPCC, 2014. PDF

Patent applications

First-listed inventor on all five patent applications listed below.

1. Method, Device, and Medium for Traffic Allocation.
   U.S. application 19/844,496. Filed October 2026.

2. Method, Device, and Medium for Traffic Routing.
   U.S. application 19/533,002. Filed February 2026.

3. Method, Device, and Medium for Security Management.
   U.S. application 19/532,711. Filed February 2026.

4. Method, Device, and Medium for Cause Analysis.
   U.S. application 19/529,988. Filed February 2026.

5. Fail-slow Disk Detection Method, Device, Computer Equipment and Computer Readable Storage Medium.
   Chinese application CN202410263229. Filed March 2024; published September 2025 as CN120610857A.
   Google Patents

Peer review and professional service

IEEE Transactions on Software Engineering 2026
Reviewer, TSE.

IEEE Transactions on Dependable and Secure Computing 2026
Reviewer, TDSC.

IEEE International Conference on Network Softwarization 2026
Reviewer, NetSoft.

ACM Transactions on Architecture and Code Optimization 2025
Reviewer, TACO. Completed reviews for three manuscripts.

Software: Practice and Experience 2020
Sub-reviewer, SPE.

IEEE International Symposium on Reliable Distributed Systems 2018
Sub-reviewer, SRDS.

IEEE International Conference on Parallel and Distributed Systems 2017
Sub-reviewer, ICPADS. Big Data and Cloud Computing track.

Teaching

Teaching Assistant, The Chinese University of Hong Kong

Introduction to Cloud Computing and Storage Fall 2017, 2018, 2019
CSCI 4180 / ESTR 4106

Data Communication and Computer Networks Spring 2018, 2019
CSCI 4430 / ESTR 4120

Teaching Assistant, University of Science and Technology of China

Computer Organization Spring 2015, 2016
01145

Selected honors and awards

Excellent Teaching Assistant Award, CUHK 2019
CSCI 4430 / ESTR 4120: Data Communication and Computer Networks.

CUHK Postgraduate Studentship 2017–2021

First Class Academic Scholarship, USTC 2014–2017

Guanghua Scholarship, USTC 2016

Third Prize, National Graduate Mathematical Contest in Modeling 2014

Meritorious Winner, Mathematical Contest in Modeling 2014

Outstanding Graduate, South China Normal University 2014

National Endeavor Scholarship 2011–2013