

Yu-Chen CHENG

<https://rudeigerc.dev> · rudeigerc@gmail.com

EDUCATION

Shanghai Jiao Tong University

M.Eng. in Software Engineering

Thesis: *Distributed Data Processing Job Scheduling in Multi-Tenant Shared Clusters*

Shanghai, China

Sep. 2019 - Mar. 2023

Shanghai Jiao Tong University

B.Eng. in Software Engineering

Shanghai, China

Sep. 2015 - Jun. 2019

EXPERIENCE

Heywhale

Software Engineer

Shanghai, China

May. 2025 - Present

Kubernetes Go LLM

WizardQuant

AI Platform Development Engineer

Shanghai, China

May. 2023 - Sep. 2024

- **AI Platform.** Participated in the redesign and refactor of internal MLOps platform upon Kubernetes, including the refactor of Python SDK and CLI, the redesign of unified Job models, the implementation of microservices via Protobuf and gRPC, and the development of dashboard via React and Vue.
- **DevOps Pipeline.** Designed and built efficient Python and Go DevOps pipelines upon Tekton to accelerate the developing and testing cycles.
- **Large Language Models.** Designed and conducted a series of experiments on large language models inference backends, focusing on metrics including Time to First Token (TTFT), Time Per Output Token (TPOT) and end-to-end throughput.
- **Technical Culture.** **a)** Established team-wide modern Python development standards with PDM. **b)** Introduced Helm templates for hosting static sites on Kubernetes with integrations of Tekton pipelines and OAuth Proxy. **c)** Built a unified documentation generation tool based on VitePress.
- **Mentorship.** Mentored **a)** the development and integration of the Dex identity provider with LDAP, providing an authentication and authorization middleware for internal systems. **b)** the benchmarking framework for Kubernetes schedulers, providing a unified interface for benchmarking and comparing different schedulers.

Kubernetes Go Python Ray

Ant Group

Site Reliability Engineer Intern

Shanghai, China

Jun. 2021 - Sep. 2021

- **Federation Observability Platform.** Participated in the design and development of the internal multi-clusters observability platform.
- **Identity Provider Connector.** Led the development of the Dex identity provider plugin connecting to the internal identity and access management system, providing OIDC services for internal systems such as Argo CD and Grafana.

Kubernetes Go React

The Linux Foundation

Mentee of LFX Mentorship Program, CNCF - Volcano

Remote

Feb. 2021 - May. 2021

- **System Stability Enhancement.** Built an efficient testing framework based on Kubernetes and imported unit tests and E2E tests to cover more classical scenarios for Volcano, a Kubernetes native batch scheduling system for compute-intensive workloads.

Kubernetes Go

JD.com

Software Development Engineer Intern

Shanghai, China

Jun. 2018 - Sep. 2018

- **JingZao.** Refactored module JingZao with React Native and introduced several sections in Objective-C for JD.com mobile application.

iOS Objective-C React Native

SERVICES

- Mentor of Google Summer of Code 2025, Kubeflow, **Project 10: Support Volcano Scheduler in Kubeflow Trainer**

PROGRAMMING SKILLS

Languages: Go, Python, TypeScript, JavaScript, C++

Skills and Expertise: Resource Management, Scheduling

Tools and Frameworks: Kubernetes, Ray, React, Vue

LICENSES AND CERTIFICATIONS

Japanese-Language Proficiency Test Level N1

2017

HONORS AND AWARDS

The Second Prize in Taiwan Student Scholarship, Ministry of Education, P.R.C.

Dec. 2018, Mar. 2021 and Oct. 2021

The Third Prize in Taiwan Student Scholarship, Ministry of Education, P.R.C.

Dec. 2017

MISCELLANEOUS

Blog: <https://rudeigerc.dev>

GitHub: <https://github.com/rudeigerc>

LinkedIn: <https://www.linkedin.com/in/rudeigerc/>

Languages: Chinese (Native proficiency), English (Professional working proficiency), Japanese (Limited working proficiency)

Selected Open Source Contributions: Member of **Tekton** and **Tensorchord**. `tektoncd/cli`, `tektoncd/triggers`, `tektoncd/dashboard`, `volcano-sh/volcano`, `tensorchord/envd`, `ray-project/kuberay`

NOTE

I served in military training service in Taiwan from October 2024 to February 2025.

Last Updated: 2025-05-26