Edward Xu

Senior Backend Engineer | Distributed Systems & Cloud Infrastructure

github.com/LindaSummer

Profile

Senior backend engineer focused on distributed systems and cloud infrastructure. Owned a globally deployed Go/C++ threat intelligence service, from architecture and API design through AWS and Kubernetes operations, scaling, observability, and incident recovery. Apache Kvrocks committer and CPython free-threading contributor.

Selected public work: Kvrocks community work CPython performance fix PyCon talk

Professional Experience

Citrix — Senior Software Engineer May 2025 – Present

Stateful telemetry processing and workload isolation

- Designed and deployed an Azure Synapse/PySpark pipeline for stateful telemetry processing, isolating batch computation from Azure Data Explorer query workloads.
- Implemented session reconstruction and deduplication from batched telemetry events, persisting derived state in partitioned Parquet.
- Corrected derived session data as late telemetry arrived, recalculating only affected partitions rather than reprocessing the full dataset.

BATechWorks — Senior Cloud Developer Oct 2024 – May 2025

Python execution orchestration and virtual ECU testing

- Built Python orchestration for VM startup and containerized electronic control unit (ECU) tests, integrating a QEMU simulator supplied by another team and propagating execution/trace context across VM, container, and simulator boundaries.
- Enabled cross-component failure diagnosis with OpenTelemetry and TTY access; provisioned AWS and Kubernetes infrastructure with Terraform and Pulumi, supporting virtual tests with less reliance on physical devices.

Trend Micro — Senior Software Engineer Jul 2020 – Oct 2024

Threat intelligence services, cloud infrastructure, and reliability

- Owned a Go/C++ threat intelligence service from requirements and architecture through API design, testing, and production operations.
- Separated query and management workloads so reads could scale independently while supporting live pattern updates.
- Used benchmarks and production metrics to guide caching, data-access, and data-store choices, balancing latency, data freshness, and database load.
- Designed failure handling and degraded operating modes to preserve critical service behavior during dependency failures.
- **Platform & operations:**
  - Designed and operated AWS/EKS infrastructure with Terraform and CloudFormation: networking, upgrades, workload/node scaling, and graceful shutdown.
  - Built Prometheus/Grafana monitoring and log pipelines for on-call diagnosis; planned shared or dedicated database and middleware capacity.
  - Matched compute and storage lifecycle choices to workload needs and cloud cost constraints.
  - Partnered with QA on CI/CD workflows and golden-image builds, using load tests, correctness probes, and failure-case coverage to validate service behavior.

Earlier Career (2017–2020): Luculent & PDStars — C++ desktop and embedded development.

Open-Source Engineering & Community

Apache Kvrocks — Contributor; Committer since May 2025 Jun 2024 – Present

Redis-compatible key-value database built on RocksDB

- Led t-digest feature planning for approximate quantile queries: defined the roadmap, decomposed tasks, coordinated assignments, tracked progress, guided contributors, and reviewed changes.
- Designed RocksDB-backed storage encoding for metadata, buffered observations, and centroids; implemented core t-digest operations and quantile calculations.
- Adapted Apache Arrow and RedisBloom implementations to Kvrocks' storage model; implemented cumulative distribution function (CDF) calculations with unit and integration tests for behavior and edge cases.
- Served as release manager for Kvrocks 2.13.0, proposing the release window, cherry-picking commits, and coordinating mailing-list voting through the Apache release process.

CPython — Contributor Dec 2024 – Present

Free-threading correctness and performance

- Investigate, reproduce, and fix free-threading correctness and performance issues, contributing fixes and regression tests upstream.
- **Performance highlight:** Fixed poor multithreaded scaling of dataclass initialization by enabling deferred reference counting for functions assigned to types; added a scaling benchmark and landed the fix upstream.

Speaking

PyCon China 2026 — Speaker Sep 2026

Contributing to CPython through Free-Threading

- Presented a practical contribution workflow using reproducible concurrency bugs, Python-to-C source analysis, ownership and synchronization rules, and focused validation.

Technical Skills

Languages: Go, C++, Python

Backend & distributed systems: Service/API design, caching, concurrency, failure recovery; Redis, Aurora MySQL

Cloud & IaC: Linux, AWS, Kubernetes/EKS, Terraform, CloudFormation, Pulumi

Delivery & reliability: CI/CD, load testing, graceful shutdown, capacity planning

Data processing: PySpark, Azure Synapse, Azure Data Explorer, Parquet

Observability: Metrics, logs, traces; Prometheus, Grafana, OpenTelemetry

Diagnostics: Concurrency debugging, profiling, tracing, reproducers, benchmarks

Education

Nanjing University of Aeronautics and Astronautics Sep 2013 – Jun 2017

Bachelor of Engineering in Electrical Engineering