Yiqiao Qiu

yiqiaoqiu@hotmail.com | 341-732-9006 | www.linkedin.com/in/yiqiao-qiu-12013723a/

Massive-scale production distributed system development and operation experiences in AWS Datacenter network fabrics routing control plane, with extraordinary understanding in **Software Defined Networks (SDN) and distributed system**: CAP theory, multi-phase commit transations, and system communications behavior/control/optimization, multiple network stack protocols; understanding both large cluster networking and AI training/inference workloads and algorithms.

SKILLS

- Programming Languages: Python, Rust, C/C++, Java, Kotlin, shell, MySQL
- Frameworks: Rust Tokio Async, Rust Netlink, AWS (S3, DynamoDB, CloudFront), gRPC, Docker, SQL, PyTorch
- Networking protocols: Quagga zebra, BGP, TCP/UDP

WORK EXPERIENCE

Amazon Web Services Santa Clara, CA

Software Engineer, AWS DC Network Infra, Scalable Intent-Driven Routing (SIDR) **(Rust, Python, Tokio)** Apr 2024 - Present

- Designed and built an end-to-end automated release qualification framework for SIDR with **hierarchical concurrent workflow** engines, comprehensive protocol state tracing, and chaos fault injection (partitions, device failures, interface flaps) to validate **consensus convergence** and MPC operation correctness with rollback under extreme stress and chaotic network/software failure conditions. Integrated into CI/CD pipeline, achieving **15x speedup** in release qualification and qualifying all subsequent SIDR production releases.
- Delivered SIDR daemon logic for network cross-protocol route redistribution, including inter-process communication, message stream parsing, multi-module **asynchronous programming and OS-signaling**, and optimized message generation and MPC lazy-init in multiple SIDR modules for improved efficiency.
- Delivered SIDR protocol security enhancement through message authentication and verification mechanisms, with **thorough design and implementation for service-fabric workflows, and race-condition/failures handling**, and comprehensive end-to-end integration tests spanning controller-to-daemon intent distribution, system-level signaling, and protocol state machine transitions under various network and system conditions.

XPeng Motors San Diego, CA

Computer Vision Engineer Intern **(Python, PyTorch, DALI, ONNX)** Oct 2023 - Mar 2024

- **Training pipeline acceleration**: Accelerated model training with **NVIDIA DALI framework** with multi-processing GPU-based huge-scale image dataset online augmentation, achieved **7x speed** and **80% reduction of used CPU resource**.
- **Eye Action video classification model development**: build up entire eye action task pipeline from making model dataset, model training and optimization, and real-scene testing, merging into one backbone, with various model structures, FLOPs, and generalization ability explored, reducing on-car model system scheduling cost.
- **Data simulation**: Replenishing an object detection dataset with simulation data of rare categories. Training YOLO-X on the replenished dataset to prove the effectiveness of the pipeline.

ByteDance Shenzhen, China

Video Algorithms Engineer Intern, Real-Time Communication, Video Group **(Python, PyTorch)** Nov 2021 - Apr 2022

- Proposed novel auxiliary modules at the low level of encoder/decoder for **real-time multi-frame Super-Resolution** models to alleviate blocky effects in video compression. Improved **43%** of the PSNR gain in TikTok/Lark offline testing.
- Optimized **facial landmark detection** model in following ways: preprocessed data with facial parsing, applied weighted loss and balanced resampling to enhance robustness and added an extra branch and loss for global information. Reduced **67%** of NME Loss in offline testing. Used **unstructured model pruning** to reduce **20%** of inference time.
- Performed offline random data augmentation, and implemented a **multi-thread concurrent** read/write queue to reduce I/O during training. Reduced **40%** of training time on the massive dataset.

Amazon Seattle, WA

Software Engineer Intern, AWS CloudFront **(Java, Kotlin, AWS CloudFront)** June 2023 - Sept 2023

CloudFront Function (CF2) Tagging in CloudFront (CF) **Control Plane**

- In **micro-service** and **distributed transaction** scenario, designed and introduced a unique ID, analyzed API call-flows, multiple exceptions, and race-conditions (caused by failure and concurrent requests), designed complete customer error message, and handled with synchronous Database entry erasing calls, cleaner threads and distributed locks in **DynamoDB**.
- Introduced a new update call for CF2 in CF internal RPC Service, and a new correctness validation component for Database entry updating. Optimized 3 customer CF2 APIs latency by **25%** through reducing redundant RPC interaction calls.
- Optimized CF Distribution APIs to greatly shorten Tagging cleaner lists and reduce the RPS of cleaner API for **98.5%**, achieved **60x** cleaning speed to ensure customers experience.
- Extended wrapper package of CF to AWS Tagging service. Implemented comprehensive integration tests for CF2 Tagging APIs, covered basic and concurrent race-condition cases.
- Extended AWS CloudFormation interface for CF2 tagging, used **asynchronous process, callbacks** and **Factory pattern**.

EDUCATION

University of California, San Diego La Jolla, CA

Master of Science in Computer Science and Engineering GPA: 3.93/4.0 Sept 2022 – Mar 2024

Sun Yat-sen University Guangzhou, China

Bachelor of Engineering in Computer Science Major GPA: 3.94/4.0 (top 10%) GPA: 3.8/4.0 Sept 2018 - Jun 2022