Yun-Chi Pang

Santa Clara, CA | 669-274-6138 | yunchipang@gmail.com | linkedin.com/in/yunchipang | github.com/yunchipang

Education

Northeastern University, MSc in Computer Science	09/2022–05/2025
The Chinese University of Hong Kong, BBA in Business Analytics	09/2017–12/2021

Technical Skills

Languages & Backend: Python, Java, Rust, SQL, Django, FastAPI, Spring Boot, Node.js
Distributed Systems & Data Infra: Kafka, Airflow, Spark, Flink, Iceberg, HDFS, ClickHouse
Databases: MySQL, PostgreSQL, MongoDB
Cloud & DevOps: AWS (EC2, S3), GCP, Docker, Kubernetes, Jenkins, GitHub Actions
Testing & Developer Tools: Git, Pytest, JUnit, Postman, Swagger/OpenAPI

Open-Source

Apache Kafka *15+ PRs* *Event streaming platform* — 03/2025–
- Migrated core Scala modules to Java, removing tech debt and reducing contributor onboarding time by 50%
- Optimized config API, reducing CPU usage by 20% and memory allocations by 60% through algorithmic tuning and lazy evaluation

Apache Hudi *10+ PRs* *Data lake platform* — 07/2025–
- Added file-slice partitioning APIs enabling 8× faster incremental processing across distributed ETL workers
- Introduced column-stats data skipping, reducing scanned data by 90% and accelerating selective queries by 10×
- Reduced query latency by up to 100× via automatic timestamp-to-partition filter transformation, eliminating 95% of unnecessary scans

Apache Gravitino *20+ PRs* *Federated metadata lake* — 04/2025–
- Cut authorization overhead by 60% through permission pushdown to catalogs
- Reduced integration test boilerplate by 60% through infra refactoring and ID autogeneration

Additional contributions across major data-infra ecosystems (Airflow, ClickHouse, OpenLineage, Sedona, OpenDAL, Ruff), focusing on reliability, performance, and developer experience.

Professional Experience

Software Engineer Intern @ OurDate — 07/2023–10/2023
AI-powered growth-stage social platform focused on date-planning
- Boosted system throughput by 25% by developing scalable REST APIs (Django, MySQL) powering core date-planning features—reduced latency for 100+ daily active users
- Suppressed post-release defects by 20% through systematic peer code reviews and mentoring junior engineers

Data Engineer @ Dashmote — 01/2022–06/2022
Series B SaaS data analytics company serving Fortune 500 clients across 50+ global markets
- Built and optimized AWS pipelines processing 10M+ records/day, cutting runtime by 88%
- Designed scalable, multi-region scraping framework (Python, Scrapy) ingesting 1M+ menu records/day across global markets

Ops Intern @ ShopBack — 09/2020–07/2021
Regional e-commerce scale-up with 35M+ users across Asia-Pacific
- Increased ops efficiency by 90% by automating manual workflows with Python tools
- Cut 98% of support tickets through Selenium-based tracking automation—drove 1% weekly partner growth