

# YUN-CHI PANG

Santa Clara, CA | 669-274-6138 | [yunchipang@gmail.com](mailto:yunchipang@gmail.com) | [linkedin.com/in/yunchipang](https://linkedin.com/in/yunchipang) | [github.com/yunchipang](https://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  
Databases & Storage: MySQL, PostgreSQL, MongoDB, Hive, Iceberg, HDFS, ClickHouse  
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 used by 80%+ of Fortune 100 03/2025–

- **Eliminated tech debt across 200K+ deployments** by migrating core Scala modules to Java, reducing contributor onboarding time by 50% and simplifying long-term maintenance
- **Slashed CPU overhead by 20% and memory allocations by 60%** in config API through algorithmic optimization and OOP refactoring—leveraged polymorphic overloading and stream-based lazy evaluation

Apache Hudi  10+ PRs Data lake platform used by Uber, Amazon, ByteDance 07/2025–

- **Enabled up to 8x faster incremental query processing** by implementing file slice partitioning APIs, distributing time-bounded data changes across workers for parallel ETL pipelines
- **Accelerated query performance by up to 10x** through column statistics-based data skipping—file-level pruning reduces data scanned by 90% for selective queries on large tables
- **Minimized query latency by up to 100x** via automatic timestamp-to-partition filter transformation—partition pruning eliminates 95% of irrelevant scans for time-range queries

Apache Gravitino  20+ PRs Federated metadata lake used by Uber, Pinterest, Xiaomi 04/2025–

- **Cut authorization overhead by 60%** through permission pushdown—delegated auth to underlying catalogs (relational DBs, warehouses, lakes), removing redundant API-layer checks
- **Streamlined test authoring by 3x** via infrastructure refactoring—auto-generated test IDs and localized dependencies, slashing boilerplate code by 60%

## PROFESSIONAL EXPERIENCE

Software Engineer Intern @ OurDate 07/2023–10/2023

AI-powered growth-stage social platform focused on event-driven matchmaking

- **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

- **Compressed pipeline runtime by 88%** by building and optimizing automated AWS data pipelines processing 10M+ records daily—improved data freshness for client dashboards
- **Expanded product coverage by 3x** through distributed web scraping framework (Python, Scrapy) ingesting 1M+ restaurant menu records daily 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