

SUMMARY

Senior Backend Engineer with 6 years building highly-available, scalable cloud systems processing millions of daily transactions. Grew from individual contributor to technical owner of 4 microservices. Led architecture design for 10+ Technical Designs across microservices, ML systems, and real-time pipelines. Highly adaptive to emerging technologies (AI agents, LLMs, modern tooling). AWS Certified Solutions Architect. Active technical blogger at cuongkane.com.

EDUCATION

- **Master of Science in Computer Science** Sept. 2021 – Sept. 2025
HCM University of Science Ho Chi Minh City, Vietnam
- **Honor Program - Bachelor of Science in Information Technology** Sept. 2017 – Sept. 2021
HCM University of Science Ho Chi Minh City, Vietnam
- **Machine Learning Course (Facebook-sponsored)** Aug. 2019 – Dec. 2019
Coder School Ho Chi Minh City, Vietnam

EXPERIENCE

- **Parcel Perform** Ho Chi Minh City, Vietnam
Senior Backend Engineer Aug 2024 - Present
 - **System Architecture & Design:** Spearheaded 10+ architecture design reviews for ML systems, analytics platforms, and real-time pipelines. Authored cross-team TD for ML-based carrier identification across 7 microservices, reviewed and approved by CTO.
 - **Microservices & Database Splitting:** Decomposed 2 monolithic services into microservices and executed 2 database splits, improving system modularity and team autonomy. Coordinated cross-squad releases with backward-compatible APIs for zero-downtime migrations.
 - **High-Availability Cloud Systems:** Delivered mission-critical Reports & Analytics platform (Jaspersoft/ClickHouse) and Alert Management system using DDD architecture with Dagster orchestration. Accelerated chart loading through static file caching, enhancing customer dashboard experience.
 - **Real-Time Data Pipelines:** Orchestrated ML-based carrier identification pipeline spanning 7 microservices (Flink CoreJobs, Scraper, DataIO, Scheduler), processing millions of daily events. Integrated seamlessly with legacy regex flow using feature toggles and Kafka protobuf schemas, boosting parcel found rate while preventing recursive scrape loops.
 - **AI/ML Solutions:** Engineered L3 AI agents using LangChain/LangGraph with ReAct pattern and AWS Bedrock. Implemented tool-based architecture for Root Cause Analysis and automated insights generation. Reduced LLM costs by 20% through strategic model selection while maintaining output quality.
 - **Engineering Excellence:** Pioneered AI-assisted Technical Design workflow, becoming first squad adopter company-wide. Delivered 105+ tickets as sole backend engineer with 100% on-time delivery. Shipped AI features under tight deadline for CEO conference demo.
 - **Release & Deployment Strategy:** Established release protocols using canary deployments, A/B testing, and feature toggles. Introduced 4-day advance deployment booking and mandatory QA sign-off, achieving zero-downtime releases with minimal incidents.
 - **Observability & Sustainable Growth:** Established observability practice with Grafana dashboards tracking P95/P99 latencies, Kafka consumer lag, and success/failure rates. Championed dashboard-first approach enabling data-driven capacity planning and proactive debugging.
 - **Mentorship & Knowledge Sharing:** Mentored 2 engineers to promotion-readiness through architecture discussions, code reviews, and pair programming. Delivered 15+ Open Space sessions on system design, AI/LangChain, and Flink, fostering bidirectional learning culture.
 - **Engineering Leadership:** Conducted 6+ technical interviews for junior/mid-level candidates. Established runbook and postmortem culture for incident response. Championed coding standards (PEP8, TDD, Django structure) and led backlog grooming with story point re-estimation.
- **Parcel Perform** Ho Chi Minh City, Vietnam
Backend Engineer Jul 2022 - Aug 2024

- **Scalable API Development:** Launched Public Booking API with OpenAPI spec serving multi-carrier integrations (DHL, MyParcel, Shippo). Implemented S3 pre-signed URLs, Kafka async processing, and Celery-based rate limiting to protect third-party APIs.
- **Extensible Integration Architecture:** Architected base booking integration abstraction supporting 50+ carrier integrations. Validated requirements through practical usage testing and Product collaboration to build flexible, extensible foundation.
- **High-Volume Data Pipelines:** Scaled data pipeline to 5M+ daily events across scheduler, scraper, data I/O, and Apache Flink services. Maintained sub-second latency for mission-critical shipment tracking.
- **Cross-Functional Collaboration:** Authored 4+ Technical Designs collaborating with Product and Design through HLD meetings. Applied ubiquitous language (DDD) to align domains and negotiated deliverables presenting trade-offs on resources and technical constraints.
- **Email Extraction Platform:** Engineered pattern-based email extraction engine (Regex/XPath/custom formatters) with Kafka CDC pipeline. System running 3+ years in production processing shipment notification emails.
- **Architectural Decision Records:** Documented architectural decisions with rejected alternatives and trade-off analysis. Evaluated 4 approaches for email processing; selected app-based CDC, recording rationale for team alignment.
- **Sync-to-Async Reliability Pattern:** Architected fault-tolerant Booking API spanning 5+ microservices with sync-to-async fallback pattern. Implemented Kafka-based async processing ensuring booking completion despite component failures.
- **Celery Pipeline Architecture:** Orchestrated Collection Point data pipeline using Celery canvas patterns (chain, chord/group) with optimized worker configuration. Parallelized 8K+ postal codes into chunked sub-tasks. Selected Celery Beat for auto-retry and fault tolerance.
- **System Reliability & Incident Response:** Diagnosed and resolved critical production incidents: BFCM traffic spikes (10x load), Redis cluster hotspots. Built recovery script in 24 hours restoring 1.5M records across multiple services for enterprise customer.
- **Large-Scale Data Migrations:** Executed production data migrations handling SQL timeout constraints, autovacuum optimization, and strategic indexing. Implemented incremental update patterns ensuring data consistency across distributed systems.
- **Testing & Developer Experience:** Authored company-wide testing best practices guide (AAA pattern, black-box testing, flaky test prevention). Drove tool adoption: Poetry, Ruff/Pyright, Django Ninja. Created onboarding documentation enabling 5 engineers to ramp up faster.

• Cinnamon AI Labs

Ho Chi Minh City, Vietnam

Software Engineer

Aug 2020 - Jul 2022

- **Enterprise Client Delivery:** Shipped production systems for 4 Fortune 500 clients (Deloitte, Bridgestone, Daikin, Sumitomo) in document processing and AI automation domains.
- **Multi-Framework Expertise:** Mastered 3 web frameworks (Flask, FastAPI, Django) adapting to varied client requirements. Designed RESTful APIs using Repository-Services pattern and SOLID principles.
- **Multi-Cloud Infrastructure:** Provisioned and maintained services across AWS, Azure, and GCP for enterprise clients. Containerized applications with Docker enabling consistent cross-cloud deployments.
- **Cross-Team Collaboration:** Collaborated with AI researchers and product managers to translate ML models into production-ready solutions meeting client SLAs.

PROJECTS

- **Codlis - AI Code Visualization (Master's Thesis, H2 2025) (codlis.com):** Multi-agent LLM system for interactive code comprehension. Architected 5-agent pipeline (Scanner, Analyzer, Architect, Illustrator, Proofreader) with Knowledge Graph and ReAct reasoning. User study: 2.9x task success rate, 29% cognitive load reduction vs baseline (p<.001). Stack: Django/Celery, Next.js, LangChain.
- **SweatCharge - Sports Club Management (H1 2025) (sweatcharge.com):** Full-stack member management featuring role-based SSO (owner/manager/player permissions), fuzzy name matching for player association, and automated cost splitting with billing. React/TypeScript with production users.

CERTIFICATIONS

- **AWS Certified Solutions Architect - Associate**
Valid: Jan 2025 – Jan 2028

Amazon Web Services

AWARDS & RECOGNITION

- **Parcel Perform (2022-Present)**: Member of the Month (3x), Best Presentation Award, AVA AI Award for AI adoption, Professional Style Award.
- **Cinnamon AI Labs (2020-2022)**: Employee Recognition Award for outstanding technical contributions.

TECHNICAL SKILLS

- **Architecture**: High-Availability Systems, Microservices, DDD, System Design, RESTful APIs, TDD, Idempotency, Outbox Pattern, Feature Toggles, SSO/OAuth
- **Big Data & Pipeline**: Apache Flink, Kafka, Celery (canvas: chain/chord/group), Dagster, RabbitMQ
- **Infrastructure & Observability**: AWS (EC2, S3, Bedrock), Docker, Grafana, Sentry, Alerting Systems, CI/CD
- **AI/ML**: LangChain, LangGraph, LangSmith, AI Agents, Prompt Engineering
- **Languages**: Python, Scala, Java, TypeScript, SQL
- **Databases & Data Stores**: PostgreSQL, Redis, ClickHouse
- **Web Frameworks**: Django, Django Ninja, FastAPI, Pydantic, Next.js/React (side projects)

LANGUAGES

- **English**: B2 Upper Intermediate – Aptis ESOL (British Council) 176/200
- **Vietnamese**: Native