

陳建豪 Jimmy Chen

DEVOPS / SRE / AI ENGINEER

jimc1682000@gmail.com • github.com/jimc1682000 • linkedin.com/in/594jimmychen • jimc1682000.blogspot.com

• Taipei / Tainan

A 13-year IT career spanning systems engineering, backend development, and DevOps / SRE. Starting in network and server operations, moving through Java web backend development, and in recent years specializing in large-scale cloud platform operations and automation. Experienced in AWS at scale, Kubernetes, observability, CI/CD, and information security (CEH), and bringing AI / agentic workflows into operations. Full history below.

CORE SKILLS

13 years across ops — from systems engineering to AI agents in production

Core strengths : Large-scale AWS • Kubernetes & automation • AI agent engineering

- **AI / Agentic** : Claude Code 、 Skills / MCP 、 CI/CD guardrails 、 Agentic workflows 、 RAG / knowledge base 、 AWS Bedrock
- **Cloud / Platform** : AWS (deep) 、 EKS 、 Kubernetes 、 RDS / Aurora 、 DynamoDB 、 Elemental 、 SQS 、 Docker 、 ECS 、 Lambda 、 S3 、 Secrets Manager 、 Bedrock 、 Personalize 、 DevOps Guru 、 GCP
- **Reliability / SRE** : Incident response 、 DR / resilience drills 、 Readiness review 、 High-concurrency tuning
- **IaC / CI-CD** : Terraform 、 Ansible 、 Chef 、 Gitea 、 Jenkins 、 GitLab CI/CD 、 GitHub
- **Observability** : Prometheus 、 Grafana 、 ELK 、 Filebeat 、 CloudWatch 、 DevOps Guru
- **Languages / Backend** : Java 、 Python 、 Go 、 JavaScript 、 Bash 、 Spring 、 Django 、 RESTful
- **Databases** : PostgreSQL 、 Aurora 、 MySQL 、 Redis 、 MongoDB 、 DynamoDB 、 MSSQL
- **Messaging / Streaming** : Kafka 、 RabbitMQ 、 ActiveMQ
- **Systems / Virtualisation / Security** : Linux 、 RHCE (cert) 、 CEH (cert) 、 KVM / QEMU 、 VMware 、 Vagrant 、 HAProxy 、 NFS 、 yum / RPM 、 pfSense / BIND 、 TCP/IP 、 iptables
- **Testing / Load** : JMeter 、 Locust

IMPACT HIGHLIGHTS

- Sole owner of the GitLab → Gitea migration: hundreds of repos, the whole engineering org, < 2h off-peak cutover
- Cloud cost optimization: EC2 RIs / Fargate Spot ~20% each, MediaLive Reservations ~50%
- Led the team coding-agent toolchain, shipping 5 shared skills (adopted org-wide in DevOps, bw rolling out to the Tech team)
- Built a multi-channel event-driven FAST monitoring platform (CloudWatch → SNS → Lambda → Slack / LINE, Terraform IaC)
- Led the Airflow 1.x → 2.0 upgrade (UAT + Production, carrying many DAGs)

AI / CROSS-DOMAIN TECHNICAL PROFILE

- **Agent-native engineering & AIOps:** treat AI as a co-developer — converging ad-hoc needs into a reusable coding-agent skill ecosystem and safety guardrails (CI auto-gates, context / secrets boundaries) rather than one-off scripts; and apply AI to operations — DRM / WAF / CVE audits and GCP permission governance, plus collaborating on the multilingual-assistant translation-model lifecycle (GPT-4-turbo → GPT-5.4, UAT → Prod, prod GPU root-cause)
- **Self-hosted AI platform & open-weight LLM inference:** self-host and operate an AI platform (OpenClaw on VPS) running production agentic workflows; evaluated larger open-weight MoE models (gemma-4-26B-A4B-it, Qwen3.6-35B-A3B) on AWS GPU EC2 as self-hosting candidates; separately ran query expansion + an LLM-as-judge eval pipeline via Ollama on no-GPU hardware (model selection MoE 35B → 8B dense @ NPU, with inference-infra fixes — streaming cancel, retry warmup, output budget)
- **Knowledge & methodology:** built a raw → digest → semantic-search / NotebookLM pipeline integrated with Jira / Notion, turning personal PoCs and lessons into queryable shared team memory (cutting repeated trial-and-error) and open-sourced as ai-kb; exploring Harness → Loop (from “humans design the harness, agents collaborate inside” to “humans design the loop, focus on the verification closure”), and shared HITL practices with the Tech team (“A Day with My AI Wife”)
- **Business foundation × engineering judgment:** apply a management lens (BBA training) to ops and team decisions — used data and cost / ROI reasoning to push back on a low-value proposal, redesigning it as a forcing-function escalation (span of control, bounded rationality) that routes out-of-scope problems to whoever can actually solve them; extend the same thinking as an analogy for AI-agent governance — guardrails ≈ accountability controls, scoped authorization ≈ delegation, harness / loop ≈ organizational design
- **Cloud scale & reliability (SRE):** large-scale AWS operations and multi-cloud governance; high-concurrency architecture (RDS Proxy connection pooling, SQS high-volume order smoothing); reliability engineering (multiple DR / resilience drills at TrendMicro, E2E monitoring, service readiness reviews); observability (Prometheus / Grafana / ELK / CloudWatch); Aurora Read Replica + RDS Proxy HA; Linux-level cgroup / namespace, FCOS ignition
- **Virtualisation & self-built platforms:** deployed a full OKD 4.4 (OpenShift upstream) cluster on two physical KVM/QEMU hosts, covering HAProxy / BIND / NFS / pfSense end-to-end; earlier built enterprise VMware vSphere HA environments and ran P2V migrations

EXPERIENCE

Staff DevOps Engineer • CATCHPLAY

2024.04 – Present

- **AI collaboration & DevOps toolchain:** led the coding-agent toolchain (command → plugin → skill), shipping reusable knowledge-base and account & secrets management skills; co-built an EKS reference environment with guardrails (pre-commit secret scan, PR gates, layered secret scanning); member of the ai-squad (a Tech-Lead-formed AI initiative) and mentored a transferring colleague; shared HITL practices and the ai-kb methodology publicly

- **Streaming media & FAST channel:** built a multi-channel event-driven FAST monitoring platform (CloudWatch → SNS → Lambda → Slack / LINE, alert heatmaps, automated weekly reports, Terraform IaC); integrated AWS Elemental (MediaLive / MediaPackage / MediaTailor), DRM, Global Accelerator, SPEKE, IVS, and Elemental Link; managed MediaLive reservation purchasing and cost (~50% savings)
- **Platform modernization & service migration:** sole owner of the GitLab → Gitea migration (hundreds of repos, the whole engineering org, < 2h off-peak cutover); built Vaultwarden and rolled out Bitwarden; evaluated and adopted Bruno to replace Postman; fixed the Akamai log server and completed LDS → DataStream 2 IaC; built Serverless Lambda Terraform modules and Jenkins automation
- **Database, cache & deployment automation:** Aurora read replicas, VCS-platform database MySQL / PostgreSQL parallel operation, RDS Proxy to absorb bursts of thousands of simultaneous connections, Redis 5 → Valkey 7 upgrade (cost + EOL, some legacy systems retained); led the Airflow 1.x → 2.0 upgrade (UAT + Production, carrying many DAGs); developed Jenkins deployment automation in Go (geo-block workflow); designed high-traffic order processing with SQS
- **Cloud governance, security & client projects:** GCP permission tightening and API-key cleanup; client-project environments and CDN network allowlists (Akamai / CloudFront); AI-assisted DRM / WAF / CVE audits; credential cleanup (Secrets Manager); audited and cleaned up idle EC2 snapshots; introduced AWS Personalize, DynamoDB, Secrets Manager, and Bedrock
- **System integration & architecture:** built S3 / NAS hybrid storage for video content and S3 lifecycle auto-cleanup; developed Akamai IaC for automated CDN cache management; built a short-link service; optimized portal deployment (build once, deploy anywhere); built a DynamoDB monitoring dashboard and deployed DevOps Guru
- Tools: AWS (Elemental / EKS / Aurora / RDS Proxy / DynamoDB / SQS / IVS / Personalize / Bedrock / Secrets Manager / DevOps Guru), GCP, Airflow, Gitea, Vaultwarden, Akamai DataStream, Claude Code / skills / MCP, Bruno, Terraform, Kubernetes, Go, PostgreSQL

Senior SRE Engineer • TrendMicro

2021.09 – 2024.03

- Ran regular disaster-recovery drills (DR Drill, every six months to quarterly, 5–6+ over tenure) and maintenance windows across multiple core services; handled incident response
- Optimized cloud cost by identifying and removing unused / low-utilization resources, tuning data transfer, and adjusting logging
- Strengthened operations: Opsworks → ASG / SSM migration, end-to-end (E2E) monitoring, and a service readiness review process
- Tools: AWS (Opsworks / ASG / SSM / CloudWatch), Chef

Senior DevOps Engineer • GoFreight

2020.10 – 2021.09

- Containerized dev / stage environments (replacing VMs); ran unit tests in the CI/CD pipeline, improving pipeline performance and reliability
- Built Kubernetes (EKS) for high availability and rolling upgrades
- Built an ELK stack for centralized log management; deployed Filebeat to collect service logs
- Right-sized EC2 instances + RIs to reduce cloud cost
- Upgraded MySQL 5.6 to 8.0, upgraded the DMS engine and fixed DMS tasks; upgraded UTM

- Added autoscaling Jenkins workers and integrated GitHub
- Deployed and troubleshot the GoFreight FMS system for customers
- Tools: Ansible, Terraform, Docker, Node.js, Nginx, ELK, Filebeat, GitHub, Jenkins, AWS, EKS, Redis, Django, Vagrant

Senior DevOps Engineer • CATCHPLAY

2019.02 – 2020.10

- Built a one-click CI/CD pipeline for RDs to deploy across environments
- Built a stress-test environment (Locust / JMeter) to identify performance bottlenecks; tuned AWS load-balancer settings
- Replaced cron jobs with Airflow scheduling and refactored part of the ETL automation
- Cut cloud cost using EC2 Spot + RIs
- Wrote monitoring scripts with Slack alerts; used Lambda to consolidate multi-source logs
- Standardized components via containers (multi-layer images)
- Maintained Elasticsearch index lifecycle with Curator; tuned Logstash parsing rules to improve log ingestion
- Tools: Ansible, Terraform, Docker, ELK, GitLab, Jenkins, AWS, Airflow, RabbitMQ, ECS, Lambda, Locust, JMeter

DevOps Automation Engineer • Groundhog Technologies

2018.09 – 2019.02

- Led a small team (including a junior) driving deployment and operations automation
- Deployed DSP / DMP (real-time ad bidding) container images to customer Linux environments
- Removed redundant components and tuned service settings to improve performance
- Built inter-server connectivity monitoring tools for proactive anomaly detection
- Introduced Kubernetes, GitLab CI/CD, and Ansible to strengthen deployment automation
- Audited firewall settings and patched security vulnerabilities
- Tools: Docker, Kafka, Zookeeper, ELK, Grafana, Prometheus, Kubernetes, Ansible, GitLab CI/CD, GCP, MongoDB, Redis

Senior Java Web Engineer • Groundhog Technologies (Xianguan Technology)

2017.11 – 2018.09

- Maintained and built new features for the CovMo (cellphone signal optimization) system; refactored code for readability; tuned SQL queries for performance
- Tools: Java, JavaScript, GWT, Spring, Tomcat, MySQL

Senior Java Web Engineer • Bank SinoPac

2015.04 – 2017.11

- Gathered user requirements and built, from the ground up, a system digitizing the four core investment-trust workflows (analysis reports, sign-off, order placement & trading, risk control (pre / real-time / post), performance analysis, audit trail); built multiple RESTful web services with ZK, Hibernate, JMS, and GSON
- Tools: Java, ZK, Hibernate, JMS, GSON, Tomcat, ActiveMQ, MSSQL

Senior System Engineer • Gping Construction

2014.03 – 2015.04

- Hardened firewall settings and upgraded network topology (routing, VLAN, NAT)
- Built a backup server with Acronis
- Tools: Windows Server, Acronis, Network Administration

System Engineer • Gobooks Publishing

2012.09 – 2014.03

- Upgraded Windows AD (2000 → 2008) and Exchange (2000 → 2007); ran P2V migrations for legacy systems, built HA with VMware, and set up backup / DR

- Migrated Fortigate → Cyberoam UTM; deployed a Sophos antivirus console; set up WSUS for centralized patch management
- Successfully mitigated a denial-of-service (DoS) attack incident
- Tools: Windows Server, VMware, Hyper-V, Exchange, AD, WSUS, UTM, ERP

SELECTED PROJECTS

- **film-brain**
- **resume** — Jimmy Chen résumé (VitePress + Typst); base /resume/
- **a-day-with-my-ai-wife**
- **OKD4 Home Lab** — Bare-metal OKD 4.4 (OpenShift upstream) on two physical KVM/QEMU hosts; self-built pfSense / HAProxy / BIND / NFS / FCOS cluster end-to-end • [Config Files ↗](#)
- **CDN Traffic Report** — Browser-automation scraper for Akamai Control Center SPA + AWS CloudFront traffic data; also a Claude Code Plugin for AI-agent-driven weekly reporting • [GitHub ↗](#)
- **Enlighten by Item (UnionPay Mobile App Contest)** — led a 3-person backend team; shopping experience on UnionPay's payment API + a personality quiz; top-5 finalist • [Demo ↗](#)
- **Java Dojo — Refactoring Patterns** — talk on refactoring techniques
- **ZK 101** — introductory tutorial on the ZK framework

CERTIFICATIONS

- **Red Hat:** RHCE (EX200 / EX300)
- **Security / Network:** CEH (312-50) • ENSA (312-38) • CCNSP • CCNA (640-802)
- **Microsoft:** MCITP (Windows Server 2008 R2: 70-646 / 640 / 642) • MCSA • MCTS • MOS Master
- **Java:** OCPJP7 (1Z0-803 / 804) • SCJP
- **SRE / Incident:** PagerDuty Incident Responder
- **Other:** International Trade Proficiency Exam • Life Insurance Agent (LIA-ROC)

PROFESSIONAL TRAINING

- **Uniforce** “Network Systems Integration Engineer certification program” (2012.02 – 07): 516 core hours / 626 total hours, graduated 3rd in class
- **Uniforce** “EC-Council CEH ethical hacking course” (2013, 40 hours)
- **PCSCHOOL** “Android app training (incl. OCPJP)” (2013 – 14, 264 hours)
- **PCSCHOOL** “OCPJWCD (formerly SCWCD)” (2014, 48 hours)
- **Duotify** “Enterprise LLM Cluster Deployment — inferencing the full DeepSeek-R1 671B model” (2025.03, hands-on certification)

EDUCATION

- **Soochow University** — B.B.A., International Trade (2007 – 2012); VP of the Motor Vehicle Club, Aikido Club, group leader in the Young Leaders program
- **Tainan First Senior High School** — Science track (2004 – 2007); Computer Club

LANGUAGES

- Chinese (native) • Taiwanese (fluent) • English (professional working proficiency • TOEIC 845) • Japanese (basic)