



John D'Esposito

Kong API Gateway & Kong AI Gateway Architect

SUMMARY

- **Kong API Gateway & AI Gateway architect** — 20+ years in high-volume, mission-critical API management, with deep hands-on coverage of Kong Gateway Enterprise, Kong Konnect, and Kong AI Gateway — operated as code through deck, Terraform, and CI/CD.
- **Currently leading an enterprise WSO2 → Kong Konnect modernization** — for a national pharmacy-benefit-management enterprise: plugins and policies, routing, certificates, runtime groups / control planes, and developer portal — delivered wave by wave with rollback validation and no flag-day cutover.
- **Proven at production scale** — architected and automated Kong platforms handling 7B+ daily requests on AWS ECS Fargate across multiple regions and availability zones.
- **Extends control-plane discipline to LLM traffic** — through Kong AI Gateway: multi-provider inference normalization, prompt templating and guarding, risk-aware model routing, token and cost controls, and AI traffic audit logging.
- **Custom plugin developer** — in Lua, Python, and TypeScript — authorization edge cases, request and prompt validation, dual-authority token bridging, and SIEM metadata streaming.

KONG PLATFORM SKILLS

Kong products: Kong Gateway Enterprise, Kong Konnect, Kong AI Gateway / AI Proxy, Kong Mesh (Kuma / Envoy), Kong Manager, Konnect Developer Portal, deck, Insomnia

Gateway configuration: Gateway Services & Routes, plugins & policies, consumers & credentials, certificate management, runtime groups / control-plane topology, workspaces & RBAC, cross-cluster routing to backend application clusters

Deployment modes: Hybrid (control-plane / data-plane split), DB-less declarative, Konnect-managed data planes over an mTLS tunnel, AWS ECS Fargate & EKS, Kubernetes with Helm, golden Vault-bootstrapping image (EKS Pod Identity → IAM → Vault), multi-region / multi-AZ HA

Kong AI Gateway: ai-proxy & ai-proxy-advanced, ai-prompt-template, ai-prompt-guard, ai-rate-limiting-advanced; multi-provider LLM routing & inference-API normalization, prompt templating & sanitization, prompt-injection protection, risk-aware model routing, token & cost controls, token streaming, gateway-enforced AI usage policy, AI traffic audit logging

Plugins & policies: openid-connect, RS256 JWT validation, ip-restriction, rate-limiting (Redis-backed), OpenTelemetry, Prometheus, file-log, Salt Security API sensor; split plugin ownership — global policy via Terraform, service/route/consumer scope via deck

Custom plugin development: Lua & Python on the Kong PDK — authorization use cases, request & prompt validation, traffic-routing logic, AI-proxy policy enforcement, dual-authority token bridging, SIEM metadata streaming; TypeScript handler bridges

ApiOps & config-as-code: deck (dump, diff, sync, validate) with tag-scoped sync so a sync only touches managed entities; Spectral governance rulesets enforcing tagging; Terraform / Terragrunt Konnect IaC; GitHub Actions with policy-gated shared reusable workflows and per-environment OIDC; environment promotion, canary & rollback validation; GitOps / Flux

API security & identity: Okta OAuth2 / OIDC, signed-JWT (private_key_jwt) client credentials, local JWKS validation at the edge, RS256 JWT, mTLS, DPOP (sender-constrained tokens), Microsoft Entra SAML SSO, RBAC, rate limiting & quotas

Certificates & secrets: HashiCorp Vault KV (data-plane secrets) & Vault Transit (non-exportable signing keys); Vault-sourced data-plane↔control-plane mTLS with rotation in Vault and no key on node disk; NLB TCP-passthrough edge TLS with static CA anchors for CIAM JWKS trust; expiry tracking & automated certificate rotation

Developer Portal: Konnect Developer Portal configuration; API products & publishing (OpenAPI catalog + implementation link + auth strategy); Dynamic Client Registration with a custom DCR handler minting non-exportable RSA keys in Vault Transit for true portal self-service; consumer RBAC via Entra SAML groups

Migration & modernization: WSO2 API Manager 2.6.0 + IBM ISAM → Kong Konnect + Okta; brownfield (opaque-token bridge, consumers preserved) and greenfield (edge-validated RS256 JWT) tracks; dual-path routing plugin for side-by-side cutover; wave prep, test execution, defect triage, rollback validation

Observability & SIEM: Kong metrics to Grafana & Prometheus, OpenTelemetry, Fluent Bit, OpenSearch, CloudWatch, Dynatrace (OneAgent, distributed tracing, Davis AI — Advanced certified), Connect audit-log webhooks to Cribl with principal attribution, Google Chronicle SIEM streaming, break-glass data-plane runbook

EXPERIENCE

Prime Therapeutics

Sep 2024 – Present

Lead Kong API Gateway Engineer / Architect

Leading the enterprise WSO2 API Manager → Kong Konnect gateway modernization for a national healthcare / pharmacy benefit management (PBM) enterprise.

- **Migration leadership** — replacing end-of-life WSO2 API Manager 2.6.0 + IBM ISAM with Kong Konnect + Okta across a phased, ~12-month, side-by-side transition — wave preparation, test execution, defect triage, and rollback validation, with no flag-day cutover.
- **Dual-path routing plugin** — engineered a plugin validating each request against the correct authority (legacy ISAM opaque tokens vs. modern edge-validated JWTs), keeping existing consumers untouched while new APIs onboard to Kong — the mechanism that makes wave-by-wave migration and rollback safe.
- **Gateway configuration as code** — runtime groups / control planes (dev, test, prod, POC) declared in Terraform with no click-ops; Gateway Services & Routes managed in deck with tag-scoped sync; split plugin ownership — global policy via Terraform, scoped via deck — with Spectral lint enforcing tagging governance.
- **Certificates** — two engineered cert surfaces: Vault-sourced data-plane↔control-plane mTLS (rotate in Vault, no key on node disk) and NLB TCP-passthrough edge TLS with static CA anchors so Kong trusts Okta / CIAM JWKS; expiry tracking and automated rotation design.
- **Security & auth** — validated end-to-end signed-JWT (private_key_jwt) client credentials with local JWKS validation at the edge — removing per-request introspection from the hot path; Vault KV + Transit; Entra SAML SSO + RBAC on the Konnect org and Developer Portal; designing DPoP for sender-constrained tokens.
- **Developer portal & DCR** — configured the internal Konnect Developer Portal with OpenAPI catalog publishing and auth strategies; stood up Dynamic Client Registration with a custom TypeScript DCR handler minting non-exportable RSA keys in Vault Transit and registering private_key_jwt Okta clients — true portal self-service; consumer RBAC via Entra SAML groups.
- **Infrastructure & CI/CD** — DB-less, Konnect-managed data planes on AWS EKS over an mTLS tunnel, built from a golden Vault-bootstrapping image (EKS Pod Identity → IAM → Vault) so no static credential touches a node; GitHub Actions with policy-gated shared workflows and per-environment OIDC; break-glass runbook standing up or repairing a data plane in three commands; Redis-backed rate limiting.
- **Observability & acceptance evidence** — Konnect org audit-log webhooks wired to Cribl with principal attribution on every change; Salt Security API sensor as a Terraform-managed, per-control-plane global plugin; Dynatrace OneAgent tracing; authored a 41-document briefing, runbook, and configuration-reference library, each verified against live artifacts.

Technologies: Kong Konnect, Kong Gateway Enterprise, deck, Terraform / Terragrunt, GitHub Actions, HashiCorp Vault (KV + Transit), Okta, Microsoft Entra (SAML), AWS EKS / ECR / NLB, Helm, OpenTelemetry, Cribl, Salt Security, Dynatrace, Spectral, TypeScript, Python

United Airlines — Mobile Application

Sep 2022 – Sep 2024

Kong / ApiOps & AI Gateway Engineer

Owned the Kong platform for a mobile cloud transformation serving 7B+ daily requests.

- **Kong Konnect & Enterprise at scale** — ran Kong Gateway via Kong Konnect on AWS ECS Fargate supporting 7B+ daily mobile requests, fully automated through Terraform across multiple availability zones — a highly available, secure, performant reverse-proxy ingress to mission-critical backend APIs.
- **Kong AI Gateway introduction** — introduced Kong AI Gateway as a unified AI proxy fronting multiple third-party LLM providers, normalizing inference APIs and decoupling application teams from provider-specific request/response formats; prompt templating for guided usage and regex-based prompt protection enforced at the gateway.
- **Gateway-centralized AI guardrails** — moved AI guardrails out of application code and into the gateway control plane — centralized prompt protection, usage policy, credential and token management, and rate controls across LLM integrations, with unified AI observability and performance management.
- **deck-driven ApiOps** — managed Kong configuration declaratively with deck in Git, applied through CI/CD pipelines for reviewable, reproducible gateway changes and clean environment promotion.
- **Custom Kong plugins (Lua & Python)** — designed and deployed plugins for special authorization use cases, request validation, prompt / traffic-routing logic, and AI-proxy policy enforcement across business-unit and internal-IT requirements.

- **Deployment & validation testing** — drove environment-promotion, canary, and rollback validation for gateway and AI Gateway releases through CI/CD, with automated smoke and regression suites gating each deploy and defect triage feeding back into the pipeline.
- **API & AI observability** — continuous health monitoring across API and LLM traffic using Kong metrics, Grafana, Prometheus, Fluent Bit, OpenSearch, CloudWatch, OpenTelemetry, and Dynatrace with Davis AI for anomaly detection and root-cause isolation.
- **ApiOps strategy** — defined a 200+ API target-state ApiOps strategy covering quality, delivery speed, discoverability, security, and availability; Kong Mesh (Kuma / Envoy sidecars) for cross-region service-to-service mTLS.

Technologies: Kong Konnect, Kong Gateway Enterprise, Kong AI Gateway, Kong Mesh, deck, Lua, Python / boto3, AWS ECS Fargate, Terraform / Terragrunt, Grafana, Prometheus, OpenTelemetry, Dynatrace, GitHub Actions

Blackstone

Jun 2020 – Sep 2022

API Management and DevOps Engineer

International Alternative Asset Manager.

- **Introduced Kong API Gateway** — As a DevOps Architect and Engineer introduced API Management as a discipline to the firm. Brought in Kong API Gateway as a tool to centralize the securing, serving and governing APIs
- **API Management Architecture** — Established future state API management Topology with declarative governance policy management. Derived detailed and implemented the plan to migrate all disparate companywide APIs under a single control plane

Technologies: Python / boto3, GitHub Actions, Terraform Cloud, RunDeck, DataDog, AWS Serverless, Kong API Gateway

Goldman Sachs — Apple Card

Feb 2018 – Jun 2020

DevOps & Automation Architect

- **Banking-grade API infrastructure** — multi-account, multi-region infrastructure for customer-facing Apple Card transactions, including AWS API Gateway and KMS integrations; AMI pipelines with CIS security-compliance testing.
- **Secrets & compliance as code** — enterprise secrets management with HashiCorp Vault and Consul; automated compliance inspection against Goldman Sachs Tech Risk requirements.

Technologies: Python / boto3, Terraform, Vault, Consul, InSpec, AWS (API Gateway, Lambda, KMS, EC2, S3, RDS)

ADDITIONAL EXPERIENCE

- **Liberty Mutual Insurance** — DevOps & Automation Architect — Jun 2017 – Feb 2018 — DevOps toolchain for mission-critical AWS deployment with compliance-as-code and managed secrets.
- **GE Digital** — DevOps & Automation Engineer — Sep 2016 – Jun 2017 — automated secure configuration of the Predix Appliance to GE security standards using InSpec.
- **Earlier (12 years):** Raymond James • News Corp / Dow Jones • GE Capital • Federal Reserve Bank of New York • JPMorgan Chase • IBM

Industries: Healthcare (HIPAA) • Financial Services (PCI, SOX) • Insurance • Airlines • Government • E-commerce (GDPR)

CERTIFICATIONS, PATENTS, PUBLICATIONS & EDUCATION

Certifications: Service Mesh Certification • Advanced Dynatrace Certification

- US 8,972,569 — Remote and Real-Time Network and HTTP Monitoring with Real-Time Predictive End-User Satisfaction Indicator
- US 6,965,938 — System and Method for Clustering Servers for Performance and Load Balancing
- US 2006/0020699 — Method and Computer Program for Web Site Performance Monitoring and Testing

Publications: Creating an Effective DevOps Strategy • Creating ChatOps & Incident Response Solutions

Education: Lafayette College, Easton, PA — B.S., Electrical Engineering