

Zane Williamson

Lead Software Engineer · Platform & Infrastructure

San Francisco, CA · zane.w@hey.com · zane.srebench.com · linkedin.com/in/zanew · github.com/sepulworld

SUMMARY

Platform engineer building infrastructure since 2007, with deep Kubernetes, Terraform, and observability expertise. The work is always the same shape: take an opaque infrastructure surface, instrument it, turn that instrumentation into an API, then build the developer-facing tooling that makes it useful. High-velocity shipper — 165 merged PRs and ~2,000 commits in the last 17 months, most of it scaffolded with AI coding agents and shipped under review. Currently going deep on LLM inference performance (vLLM, CUDA, serving internals) toward a late-2026 pivot to labs and inference startups.

EXPERIENCE

Lead Software Engineer — Salesforce (Falcon TIDE)

2024 — present · San Francisco

- Designed and built a custom Terraform HTTP backend that captures state metadata *in-flight* during apply and persists it to DynamoDB — across **millions of state files**.
- Fronted it with an LLM query interface exposed as both a REST API and an MCP server, so engineers query the platform's Terraform footprint in natural language without piping state files through their terminals.
- Leading consolidation of a sprawled EKS fleet into a smaller, centralized set of clusters — sharper SLOs, less operator toil per engineering hour.

Senior Staff SRE — Varo Bank

2022 — 2024

- Owned platform reliability: deployed Grafana Loki + Tempo with the OpenTelemetry Operator for full-stack observability; established SLOs/SLIs across multiple engineering platforms.
- Migrated the org off Helm onto ArgoCD for deployment lifecycle; moved K8s auth from kube2iam to AWS client auth; enabled cluster autoscaling for reliability and cost.
- Built a Kafka outbox + Postgres → data-lake messaging pipeline; authored and maintained the team-wide engineering CLI in Go.

Senior Staff Engineer — Flexport

2021 — 2022

- Led greenfield design and rollout of the modern Kubernetes platform: Terraform AWS + Helm-provider-driven EKS, GitOps for infra and apps, Teleport-secured cluster access, a Python CLI pipeline entrypoint, and reusable Terraform modules for rapid iteration.
- Ran a deliberate pair-programming model with senior engineers to spread platform knowledge across the org.

Principal Engineer — Zillow

2015 — 2021

- Designed and open-sourced [cidr-house-rules](#) — a serverless API (Lambda + DynamoDB + API Gateway) auditing VPC CIDR, EIP, and NAT-Gateway utilization across dozens of AWS accounts, powering data-driven Terraform modules.
- Drove the Kubernetes-on-AWS rollout, the Istio sidecar deployment (distributed tracing + mTLS), and an Envoy proxy fronting Varnish.
- Lead contributor to [apache/solr-operator](#); open-sourced [hyper-kube-config](#) as a Kubernetes config-secret store.

Platform Engineering Consultant — Geniuslink

2013 — present · concurrent

- Own a multi-cloud Kubernetes platform end-to-end across EKS, Linode LKE, and DigitalOcean — ArgoCD GitOps, a Loki / Mimir / OTEL-Operator observability stack, Helmfile deploys with per-PR dynamic dev environments, Terraform across the footprint, Infisical for secrets, Teleport for access, and Fastly at the edge.
- Sustained full-team cadence alongside a full-time role: 165 merged PRs and ~1,600 commits across active repos in the last 17 months.

Earlier — Sysadmin → DevOps → Operations Engineering

2007 — 2015

- System Administrator → Sales Engineer → Senior Sysadmin (Xen / Puppet / Cobbler) → Operations Engineer → DevOps. Same shape of work throughout: make complex infrastructure legible to the people who depend on it.

SELECTED OPEN SOURCE & PRODUCTS

- **cidr-house-rules** — serverless AWS network-audit API.
- **apache/solr-operator** — lead contributor.
- **hyper-kube-config** — K8s config-secret store.
- **MAHA Healer** — health-data SaaS (lab parsing, trends).
- **TasteKeeper** — photo-to-recipe web + iOS app.
- **Garmin watch faces** — Monkey C; 1,000+ paid downloads.

TECHNICAL SKILLS

Orchestration

Kubernetes (EKS · LKE · DigitalOcean), ArgoCD, Helm / Helmfile, Istio, Envoy

Infrastructure as Code

Terraform (custom backends, modules), GitOps

Observability

Grafana Loki / Mimir / Tempo, OpenTelemetry, Prometheus, SLOs / SLIs

Cloud & Data

AWS (Lambda, DynamoDB, API Gateway, EKS), Linode, DigitalOcean, Kafka, Postgres

Languages

Go, Python, Terraform HCL, Monkey C; Java / Node instrumentation

AI / Platform tooling

LLM query interfaces, MCP servers, agentic developer workflows; Teleport, Infisical, Fastly

FOCUS

LLM inference performance — vLLM, CUDA, and serving internals — with the goal of contributing upstream and tuning large-scale inference deployments.

Live version & AI-velocity dashboard at zane.srebench.com