

Gerald Spencer

408-623-9465 | New York, NY | contact@geraldhspencer.com | linkedin.com/in/geraldspencer | geraldhspencer.com

Principal Software Engineer and technical leader working on NVIDIA's internal cloud for silicon engineering. Earlier work includes large-scale GPU infrastructure for AI research, fleet lifecycle automation, and developer platforms. 12+ years designing and operating mission-critical infrastructure.

Experience

Principal Software Engineer 2023 — Present
NVIDIA Santa Clara, CA

- Lead scheduling and fleet lifecycle work for NVIDIA's internal cloud for silicon engineering, focusing on capacity allocation, repeatable deployments, and safe changes across the global compute fleet.
- Founded and scaled an infrastructure organization from 2 to 100 engineers over 2 years, spanning platform and bare-metal automation and enabling 95%+ training goodput for large-scale AI research.
- Built a fleet remediation platform that triaged and resolved failures through 150K+ concurrent workflows across 50K+ GPUs, and standardized fleet operations across cloud providers.
- Delivered a 99.99% SLA CI/CD platform supporting 500+ engineers, including large-scale Bazel Remote Build Execution (RBE) infrastructure.

Staff Software Engineer 2022 — 2023
Cruise San Francisco, CA

- Architected an automated release attestation platform for a 6-person team, introducing policy-driven validation gates that reduced deployment cycle time by 40% while improving release safety.
- Designed a deployment governance framework enforcing rollout policies across the autonomous vehicle fleet, materially reducing production incidents and improving operational reliability.
- Built unified UI/CLI tooling to standardize release management workflows, enabling consistent, auditable deployments for 200+ engineers.

Staff Software Engineer 2019 — 2022
Skydio San Mateo, CA

- Led a company-wide infrastructure overhaul, achieving SOC 2 Type I certification within three quarters and establishing foundational security and compliance practices.
- Architected a high-scale CI platform supporting 15K builds/day across 2,500 ephemeral workers, using custom scheduling algorithms to optimize utilization and throughput.
- Migrated deployment infrastructure to Kubernetes, reducing release time from 2 hours to 15 minutes and enabling faster, more reliable production rollouts.

Director of Infrastructure 2012 — 2019
3Scan (now Strateos) San Francisco, CA

- Scaled infrastructure from 1 to 4 facilities, accelerating build throughput 13×.
- Architected multi-petabyte medical imaging platform: 24GB/s ingestion, 3.8PB storage capacity.
- Architected and operated custom high-throughput network infrastructure, establishing direct BGP peering with AWS over 40 Gbps fiber to enable petabyte-scale data transfer.

Education

California Maritime Academy
B.S. Mechanical Engineering with USCG 3rd Assistant Engineer License 2010 — 2014

Volunteer Leadership

Flux Foundation 2012-2023

- Board member for non-profit supporting large-scale art; lighting design lead for multiple temples and installations —network architecture, custom control software, and technical implementation.