

Justin Tahara

310-384-2847

✉ justintahara@gmail.com

linked in.com/justintahara

github.com/justin-tahara

Experience

Onyx AI

Jul 2025 – Present

San Francisco, CA

Founding Infrastructure Engineer

- Terraformed the existing AI infrastructure and codified changes to automate routinely performed upgrades for all customer environments reducing deployment complexities to **1 click**.
- Reduced the overall cost spend across Infrastructure, Tooling, Tokens by **80%** due to insufficient legacy code and refactoring existing code to be more efficient and performant.
- Handled all customer support including demos, deployments, debugging, office hours, on-call acting as the **Forward Deploy Engineer** for the company.

Applied Intuition

May 2022 – Jun 2025

Mountain View, CA

Tech Lead - Cloud Infrastructure

- Founded the **Security and Authentication** team to address real-world customer security requirements, leading the design and implementation of foundational security features for enterprise clients.
- Engineered solutions to reduce customer cloud costs by an average of **50%** a month, saving over **\$10 million** annually through infrastructure and application code optimization.
- Integrated technologies such as **Aurora DB**, **Intelligent Tiering**, and **Spot instances** using **Terraform**, enhancing cloud resource utilization.
- Built a **Golang-based** cloud service from scratch to track cost metrics for our customers **Kubernetes** clusters enabling precise cost-per-simulation-minute analysis and cost anomaly alerting through **Prometheus**.
- Designed and implemented **GDPR-compliant** infrastructure for the Global Accounts Service using **Route 53** and Terraform, managing the Elastic Beanstalk environments in multiple regions.
- Initiated **JIT Provisioning** for **SAML** and **OIDC** integrations, enhancing customer onboarding security and efficiency and added **SCIM** compatible features for our customers using SSO Providers.
- Standardized file access clients and functions across **AWS**, **GCP** and **Azure**, and introduced **K6** integration tests to preempt future regressions.
- Provided on-site cloud and on-prem infrastructure support to over 7 global B2B customers including GM, Scania, and VW group in Germany and Sweden, ensuring timely and effective assistance over a 2 month rotation.

Amazon

Jun 2021 – Sep 2021

Seattle, WA

Software Development Engineer Intern - Alexa

- Leveraged an internal tool to mock third-party API calls for Alexa Integration Tests, reducing test runtime by **90%** and eliminating dependencies that hindered efficiency.
- Created a **Java** utility class with a user guide to reduce the onboarding time of the mocking tool from 2 weeks to 1 day.
- Translated and developed new skills for Japanese users and tested functionality using an **Utterance Profiler**.

Western Digital

Jun 2020 – Jun 2021

Fremont, CA

Data Analyst - Wafer Fabrication

- Collaborated with the Heads Integration team to develop and maintain dashboards for tracking wafer metrics at US facilities.
- Performed data analysis using **Spotfire** to deliver actionable metrics for the R&D Data Engineering team.
- Utilized **IronPython** scripts to enhance Spotfire functionality, consolidating wafer information into a single location and improving user experience.

Education

University of California San Diego

Dec 2021

Master of Science in Computer Science with Depth in Artificial Intelligence

La Jolla, CA

University of California San Diego

Dec 2020

Bachelor of Science in Computer Science

La Jolla, CA

Technical Skills

Languages: Python, Golang, Java, C++, SQL, Javascript, Bash

Technologies: Kubernetes (K8S), AWS, GCP, Azure, Docker, Terraform, PostgreSQL, Aurora, cAdvisor, Druid, Grafana, Prometheus, Loki, Git, Elastic Beanstalk, S3, RDS, EC2, VPC, Route 53, Buildkite, Superset, SAML, OIDC, SCIM