Ricardo Herrera

ricardohe@outlook.com | Patterson, NY | (312) 972-1901 | ricardoherrera.co

SUMMARY

Senior Site Reliability Engineer with 9+ years building and operating large-scale infrastructure platforms. Expert in Kubernetes platform engineering (10+ clusters, 140 nodes) and AWS cloud architecture with proven success reducing infrastructure costs by 60% while achieving 99.99% uptime. Specialized in self-service developer platforms, comprehensive observability solutions (Prometheus, Grafana, Thanos), and infrastructure automation using Terraform and GitOps methodologies.

EXPERIENCE

Senior Site Reliability Engineer | Watchtower | Patterson, NY | September 2018 - Present

- Build and operate enterprise Kubernetes platform (10 clusters, 140 nodes, 1700+ pods) using Cluster API, serving 10 development teams with self-service infrastructure and 20+ daily deployments.
- Designed developer-facing platform with database operators (Redis, PostgreSQL) and standardized CI/CD pipelines, eliminating ticket-based provisioning and reducing infrastructure delivery from days to minutes.
- Implemented full-stack observability (Prometheus, Grafana, Loki, Thanos) with 5-year retention and multi-cluster federation, replacing legacy 2-week retention monitoring and enabling proactive incident detection.
- Automated private cloud infrastructure provisioning using Terraform and Ansible, reducing deployment time from weeks to 30-60 minutes and eliminating manual error-prone processes.
- Established Service Level Indicators/Objectives (SLI/SLO) framework achieving 99.95% availability for 1200+ VM platform with quarterly reviews and error budget-driven maintenance planning.
- Developed engineers from junior to senior responsibilities, enabling independent infrastructure provisioning via GitOps and operational autonomy across the Kubernetes platform.
- Designed centralized configuration management platform using AWX managing 500+ Linux VMs with GitOps-driven desired state configuration, eliminating configuration drift and reducing remediation time from hours to minutes.
- Architect and operate global VMware infrastructure spanning 3 datacenters across different regions, managing 1200+ VMs with distributed vCenter deployment and site-specific resource pools.
- Led incident response and on-call rotation, reducing MTTR by 30% through automated runbooks and improved observability.
- Reduced operational toil by 40% through infrastructure automation and self-service tooling, enabling team to focus on high-impact platform improvements.

Senior Site Reliability Engineer | Contract | G&G Outfitters | Remote | April 2022 - Present

- Architect and operate production Kubernetes infrastructure on AWS EKS with 99.99% uptime SLA, managing 620+ pods across 4 clusters and 4-5 zero-downtime deployments daily.
- Reduced infrastructure costs 60%+ by migrating from Expedient to AWS, implementing Karpenter autoscaling with spot instances while improving reliability.
- Scaled infrastructure to support 5x e-commerce traffic growth while maintaining $5K monthly AWS spend through cost optimization and elastic autoscaling.
- Optimized CI/CD pipelines enabling rapid deployment cycles across distributed services with Docker build caching and automation improvements, migrating from legacy Jenkins platform to modern GitHub Actions.
- Built self-service observability platform in Kubernetes using Prometheus, Grafana, and Loki, enabling development teams to independently monitor services, create dashboards, and troubleshoot issues through centralized portals in each cluster.
- Deployed Harbor registry with vulnerability scanning serving hundreds of daily pulls, preventing Docker Hub rate limits and securing supply chain.
- Implemented declarative EKS infrastructure with Terraform managing 3 production clusters, reducing cluster upgrades from 2-3 days of manual toil to 60-minute automated deployments via GitOps, eliminating upgrade-related downtime and risk.

Infrastructure Engineer | Olo | Remote | April 2017 - August 2018

- Built developer-facing infrastructure platform enabling development and QA teams to provision production-replica environments on-demand, reducing environment setup from 2-3 days to 15 minutes and eliminating infrastructure bottlenecks for 30+ developers.
- Designed Terraform-based infrastructure-as-code solution with reusable modules for multi-environment AWS deployments, implementing security isolation patterns and automated networking configuration while maintaining compliance across dev, QA, and production environments.
- Managed high-velocity CI/CD platform on TeamCity supporting 50+ daily builds and multiple production deployments, achieving 95%+ deployment success rate through automated testing gates and rollback capabilities for Olo’s restaurant ordering platform.

- Automated infrastructure operations using Ansible playbooks, eliminating 15+ hours/week of manual server configuration tasks and ensuring consistent state management across 100+ EC2 instances.
- Operated production and staging AWS environments for enterprise restaurant ordering platform serving national chains, managing automated deployments and infrastructure reliability for order processing services across development, QA, and production environments.

System Administrator | WSOL | Aurora, IL | October 2013 - April 2017

- Managed production VMware infrastructure in colocated data center, administering 5-10 ESXi hosts, vCenter, and Dell Compellent storage arrays (4TB capacity) supporting 200+ virtual machines for CMS clients.
- Operated Dell Compellent SAN infrastructure including volume provisioning, tiered storage optimization, and performance tuning, ensuring 99.9%+ uptime for production CMS workloads serving multiple client deployments.
- Led AWS cloud migration initiative driven by scalability and elasticity requirements, architecting infrastructure-as-code solution with Terraform for resource provisioning, Packer for automated AMI builds, and Chef for configuration management across multi-client environments.
- Automated server provisioning and configuration management using Chef cookbooks, reducing deployment time from hours to minutes and ensuring consistent state across 50+ instances supporting HubSpot-integrated CMS solutions.

EDUCATION

Bachelor of Computer Science | UCI | Mexico City, Mexico | 2011

SKILLS & TECHNOLOGIES

Core Competencies: Site Reliability Engineering, Platform Engineering, Infrastructure Automation, Incident Management, Capacity Planning, Toil Reduction
Languages: Go, Python, Bash
Cloud & Infrastructure: AWS (EKS, EC2, RDS, VPC), VMware vSphere
Kubernetes: Cluster API, Operators (Redis, PostgreSQL, ElasticSearch)
Infrastructure as Code: Terraform, Ansible, FluxCD, Carvel (kapp, ytt, kbld), GitOps
Observability: Prometheus, Grafana, Thanos, Loki, Tempo, Alertmanager
CI/CD: GitOps, GitHub Actions, GitLab CI
Reliability: SLI/SLO/SLA, Error Budgets, Blameless Post-mortems, On-call Rotation
OS: Linux (Ubuntu, RedHat), Windows Server
Container Registry: Harbor
Platform Engineering: Self-service tooling, developer experience (DevEx), internal portals