Rivaldo Silalahi

Site Reliability Engineer / Cloud Native & Observability

rivaldosilalahi13@gmail.com | ********** | Samosir, North Sumatera, Indonesia
linkedin.com/in/rivaldo-silalahi | github.com/IDiot29

Summary

Site Reliability Engineer with 3+ years of experience building and operating cloud-native platforms, observability stacks and AI inference infrastructure on Kubernetes (EKS, k3s), AWS and GCP. Focus areas: SRE, DevSecOps, GitOps and AI infra — designing centralized metrics/logs/traces, standardizing reliability practices and running GPU-based LLM inference with vLLM and SGLang. Strong bias toward automation (Terraform, NixOS, GitOps) and reducing operational toil for engineering teams.

Skills

SRE & Observability: Prometheus, Cortex, VictoriaMetrics, ELK, OpenTelemetry, Grafana, Loki
Cloud Native & OS: Linux, NixOS, Kubernetes (EKS, k3s, Kubespray), Helm, Flux, Argo CD, Docker, AWS, GCP
DevSecOps & Platform: Terraform, Ansible, Packer, Nix, CI/CD (Google Cloud Build, CircleCI, GitHub Actions), Git
AI Infrastructure & Perf: vLLM, SGLang, k6, Locust
Programming: Go, Python, Bash, SQL, Nix

Professional Experience

Senior Site Reliability Engineer Jul 2025 – Present
Lintasarta, Jakarta, Indonesia

- Own the design and standardization of a centralized observability platform using Cortex (metrics), ELK stack (logs) and OpenTelemetry as the common ingestion layer for services and infrastructure.
- Operate production k3s-based Kubernetes clusters dedicated to observability and shared platform components.
- Build custom exporters in Go to collect metrics from NVIDIA GB200 devices via BMC and expose them to Cortex / Prometheus for GPU fleet monitoring.
- Implement GitOps with Flux and Argo CD to manage observability configuration, dashboards and alerting rules declaratively.
- Partner with multiple product teams to roll out common SLOs, dashboards and alerting patterns, improving incident visibility and reducing ad-hoc monitoring.

Site Reliability Engineer Nov 2022 – Jun 2025
GDP Labs, Jakarta, Indonesia

- Provisioned cloud infrastructure on AWS and GCP using Terraform, Ansible and Packer, enabling reproducible environments for internal and client workloads.
- Operated Kubernetes clusters (EKS, Kubespray) and container workloads (Docker) for microservices and platform components.
- Built and maintained CI/CD pipelines with Google Cloud Build and complementary tooling, automating build, test and deployment workflows.

- Deployed and ran GPU-based LLM inference servers using vLLM and SGLang, and executed benchmark tests with k6 to evaluate model and inference server performance.
- Implemented and evolved monitoring and logging stacks (Prometheus, ELK, Grafana) to reduce MTTR and improve incident investigation.

Site Reliability Engineer Intern Jun 2022 – Aug 2022
GDP Labs, Jakarta, Indonesia

- Automated infrastructure provisioning using Terraform and Brainboard to standardize environment creation across projects.
- Built CI/CD pipelines with Google Cloud Build to decrease manual deployment effort for engineering teams.
- Deployed VictoriaMetrics as time-series storage for metrics to support high-availability monitoring.

Backend Engineer Student Feb 2022 – Jun 2022
Generasi Gigih

- Completed an intensive backend program focused on Ruby, Ruby on Rails, OOP, clean code and SOLID principles.
- Applied unit testing, TDD and basic system design to build a backend application on Ruby on Rails.

DevOps Engineer Intern Oct 2021 – Jan 2022
Finku, Jakarta, Indonesia

- Implemented CI/CD pipelines on CircleCI for backend services to streamline build and deployment.
- Performed load testing with k6 to validate the performance of critical API endpoints.
- Built Grafana dashboards on top of Elasticsearch to give engineers real-time visibility into application behavior.

Education

Bachelor of Information Systems (S1) 2024 – 2026
Binus University, Indonesia

Associate Degree (D3), Computer Technology 2019 – 2022
Del Institute of Technology, Indonesia