# Manish Khadka

Senior DevOps / SRE / Platform Engineer

Kathmandu, Nepal (UTC+5:45) · Remote or on-site
+977 9843729354 · ManishKhadka@protonmail.com
linkedin.com/in/manishkhadka · github.com/pen-pal
manishkhadka.info.np · palpen.medium.com

## Summary

Senior DevOps engineer with 8+ years in infrastructure and operations, nearly seven of them in DevOps and platform engineering roles for companies in crypto, ML video analytics and ed-tech. Builds and runs cloud infrastructure on AWS, GCP and Azure with Kubernetes, Terraform and GitOps, and treats reliability, security and cost as part of the same job. AWS Certified DevOps Engineer – Professional.

## Core Strengths

- **Reliability:** zero-downtime database and cluster upgrades; 99.99% uptime on node infrastructure; faster recovery.
- **Platform:** Kubernetes across 20+ clusters; company-wide Helm blueprint; GitOps; per-PR preview environments.
- **Security:** SOC 2 and ISO 27001 programmes at two companies; zero-trust access; scanning in every pipeline.
- **Cost:** halved cloud spend at two companies, including $24K to $12K a month at one.
- **ML and GPU:** GPU sharing that cut per-inference cost by over 60%; open-source GPU autoscaling contributions.

## Technical Skills

| | |
|---|---|
| **Cloud:** | AWS (EC2, VPC, S3, Lambda, RDS, EKS), GCP, Azure, DigitalOcean, Oracle Cloud |
| **Containers:** | Kubernetes (EKS, self-hosted, k3s), ECS, Docker, Docker Swarm, containerd, Podman, Helm, HPA, KEDA, Cluster Autoscaler |
| **Networking & access:** | Istio, Cilium, Kong API Gateway, Nginx, HAProxy, Caddy, Cloudflare (incl. Zero Trust), AWS WAF, NACLs, Route 53, Tailscale, WireGuard, Pritunl, Fortigate |
| **Infrastructure as code:** | Terraform, OpenTofu, Terragrunt, Ansible, CloudFormation, CDK, Packer, AWS Control Tower, Account Factory for Terraform |
| **CI/CD & GitOps:** | ArgoCD, Flux CD, GitHub Actions, GitLab CI, Jenkins, CircleCI, TravisCI, Bitbucket Pipelines, AWS CodePipeline, Fastlane, blue-green and canary releases |
| **Observability:** | Prometheus, Grafana (Loki, Promtail, Tempo, Alloy, Faro, OnCall, Pyroscope), ELK, Datadog, New Relic, CloudWatch, Sentry, GlitchTip, OpenReplay, ScoutAPM, Honeybadger, PagerDuty |
| **Security & compliance:** | SOC 2, ISO 27001, GDPR, zero-trust networking, least-privilege IAM, audit logging, PKI, HSM, SAST, HashiCorp Vault, AWS Secrets Manager, External Secrets Operator, Kyverno, KubeArmor, Trivy, Snyk, Checkov, Terrascan, ClamAV, BitDefender, SonarQube, AWS Private CA, AWS Inspector, GuardDuty, Security Hub, CloudTrail, KMS, AWS Config, Prowler, Wazuh (SIEM), AWS Well-Architected (Security Pillar) |
| **Databases & storage:** | PostgreSQL (PostGIS, TimescaleDB), Aurora and RDS, pgBackRest, MySQL, MongoDB and Atlas, DynamoDB, Redis, Elasticsearch, Meilisearch, InfluxDB, Milvus, Oracle, SQL Server, MinIO |
| **Messaging & media:** | RabbitMQ, Kafka, ActiveMQ, Jitsi, Flussonic Media Server |
| **Application stacks:** | Node.js, Ruby on Rails, .NET, Django, React, WordPress, Discourse, event-driven architecture |
| **Data & ML:** | GPU sharing and scale-to-zero autoscaling on Kubernetes, MCP, Snowflake, AWS Glue, Kinesis Firehose |
| **Reliability & cost:** | Incident response, autoscaling, disaster recovery, zero-downtime migrations and upgrades, load testing (k6, Gatling), chaos engineering (LitmusChaos), FinOps |
| **Operating systems:** | Linux (Ubuntu, Debian, RHEL family), VMware, Vagrant |
| **Languages:** | Go, Rust, TypeScript/JavaScript, Python, Shell |

## Experience

**Senior DevOps Engineer** Feb 2025 – Nov 2026

*ZenLedger · United States (remote)*

- Led the SOC 2 and ISO 27001 programme: coordinated teams, prepared infrastructure controls and documentation, and ran auditor and tooling selection.
- Migrated from Timescale Cloud to a self-hosted 3-node high-availability TimescaleDB cluster (Autobase, pgBackRest, 5 TB+ of backups) with minimal downtime.
- Upgraded Amazon RDS from PostgreSQL 14 to 15 with zero downtime.
- Led the move off Datadog to a self-hosted Prometheus, Grafana, Alloy, Faro, OpenReplay, GlitchTip and PagerDuty stack covering 100% of services, with alert thresholds tied to SLA targets.
- Replaced the legacy VPN with Cloudflare Zero Trust and Tailscale, giving identity-aware, least-privilege access to every environment.
- Hardened AWS validator and RPC infrastructure with least-privilege IAM, security group isolation and CloudTrail

audit logging.
- Ran disaster-recovery drills and zero-downtime migrations for production validator nodes and RPC endpoints.
- Ran rolling EKS and EC2 fleet upgrades across 8+ clusters, keeping instances on the latest patched images with zero service disruption.
- Automated provisioning and releases with Terraform, Ansible, Packer-built AMIs and GitHub Actions.
- Connected Prod, Stage, QA and Dev with VPC peering for private cross-account traffic.
- Cut monthly infrastructure spend by moving edge security from Cloudflare to AWS WAF.
- Built Python and Bash automation, provisioned with Terraform, that shuts down QA and Dev environments outside business hours.
- Brought LLM tooling into daily engineering practice: prompt-driven code review, AI-generated integration testbeds and structured-context agents for infrastructure work.

**Stack:** AWS, Kubernetes (EKS), Terraform, Ansible, Packer, ArgoCD, GitHub Actions, PostgreSQL, TimescaleDB, Redis, Prometheus, Grafana, PagerDuty, Datadog, ScoutAPM, Honeybadger, Cloudflare, Tailscale, Ruby on Rails, Go

**Senior DevOps Engineer (Contract)** Nov 2023 – Feb 2025
*CloudHero · Romania (remote)*

- Led a monolith-to-microservices migration on Kubernetes, improving scalability and modularity and cutting deployment times.
- Migrated client services from on-premises to AWS, GCP, Azure and DigitalOcean with near-zero downtime.
- Ran container workloads on ECS and EKS across distributed environments.
- Deployed self-hosted Kubernetes on GCP with Ansible and auto-scaling across 20+ clusters.
- Wrote Bash scripts that automate resource deployment on self-hosted Kubernetes clusters.
- Upgraded production EKS from v1.16 to v1.21, coordinating with stakeholders for minimal downtime.
- Wrote the central Helm chart blueprint adopted company-wide.
- Enforced admission control and policy on every client cluster with Kyverno.
- Wrote scripts that migrated thousands of ingress resources to Kong API Gateway.
- Introduced Istio service mesh with canary deployments for progressive rollouts.
- Eliminated static AWS secrets using AWS Private CA for IAM key rotation.
- Added Trivy and Snyk vulnerability scanning to every CI/CD pipeline.
- Provisioned cloud resources with Terraform for consistent, reproducible environments across platforms.
- Ran GitOps with Terraform Cloud, ArgoCD, GitHub Actions, GitLab CI, Jenkins and Bitbucket Pipelines for auditable infrastructure changes.
- Built monitoring and observability on Prometheus, Grafana, Loki and ELK that cut mean time to recovery.
- Set up on-call alerting with Grafana OnCall and PagerDuty.
- Mentored junior engineers and ran knowledge-sharing sessions on infrastructure as code for the team and clients.
- Wrote the onboarding documentation used by new team members and clients.

**Stack:** AWS, GCP, Azure, DigitalOcean, Kubernetes (EKS, self-hosted), ECS, Helm, Terraform, Ansible, Packer, ArgoCD, GitLab CI, GitHub Actions, Bitbucket Pipelines, Jenkins, Kong, Istio, Kyverno, Nginx, Prometheus, Grafana, Loki, Tempo, Pyroscope, ELK, Sentry, RabbitMQ, ActiveMQ, Redis, Elasticsearch, Meilisearch, MongoDB Atlas, PostgreSQL, Pritunl, WireGuard, OwnCloud, Discourse, Node Termination Handler, Cloudflare, Route 53, Cloud DNS, Azure DNS, .NET, Python, Node.js

**DevOps Engineer, then Team Lead** Jul 2021 – Feb 2024
*Innovatetech Pvt. Ltd. · Kathmandu, Nepal*

- Led infrastructure and backend engineering for the live-class platform behind mySecondTeacher, an ed-tech product serving millions of users through the COVID-19 demand surge.
- Promoted to Team Lead after seven months (Feb 2022); led and mentored the DevOps team.
- Introduced a weekly R&D process and code review standards.
- Delivered infrastructure and security reports to the CTO that shaped key decisions.
- Designed and built **JMW (Jitsi Meet Multi-tenant)**, a multi-tenancy layer that streams live JVB and Jicofo metrics into AWS to autoscale from custom AMIs on real demand instead of provisioned peak capacity.
- Ran self-hosted Jitsi for secure video conferencing, reducing reliance on external tools.
- Cut cloud spend 50% across a 10+ account AWS organisation using spot and reserved instances, right-sizing and automated shutdown outside peak hours.
- Standardised multi-account governance with AWS Control Tower and Terraform, reducing configuration errors.
- Implemented Terraform, CloudFormation and Ansible across multi-region AWS and bare-metal servers.
- Managed Cloudflare DNS and MongoDB Atlas configuration as code.
- Introduced blue-green and canary releases for zero-downtime deployments.
- Built central GitLab CI pipelines with shared Helm charts for Android, iOS, frontend and backend apps.
- Introduced a monorepo strategy and standardised rollback procedures for frontend and backend.
- Built per-PR preview environments for frontend teams, cutting feedback loops from days to minutes.
- Added revert options for frontend releases using GitLab CI and S3, so a bad release can be rolled back.

- Migrated EC2-hosted WordPress sites to Kubernetes with their plugins and add-ons.
- Drove SOC 2, ISO 27001 and GDPR compliance.
- Built shift-left security scanning into every pipeline (SonarQube, Snyk, Trivy, Checkov).
- Hardened workloads with distroless containers, ClamAV file scanning, VPC endpoints and Pritunl VPN.
- Designed the disaster-recovery strategy on automated AWS Backup.
- Ran EKS with HPA, KEDA, Cluster Autoscaler, Istio and External Secrets Operator.
- Built tracing and observability on Prometheus, Grafana, Loki, Tempo, Pyroscope, ELK and Sentry.
- Restructured MongoDB Atlas with fine-grained access controls.
- Set up VPC peering between MongoDB Atlas and AWS, cutting service-to-service latency.
- Built ETL pipelines that replicate production data to staging for reliable testing.
- Load-tested services with k6.
- Documented DevOps processes, reducing onboarding time.

**Stack:** AWS, Kubernetes (EKS), Helm, Terraform, CloudFormation, Ansible, Packer, GitLab CI, Fastlane, SonarQube, Snyk, Trivy, Checkov, ClamAV, Prometheus, Grafana, Loki, Tempo, Pyroscope, ELK, Sentry, GlitchTip, InfluxDB, k6, PostgreSQL, MongoDB Atlas, Redis, Elasticsearch, Meilisearch, RabbitMQ, Jitsi, Pritunl, WireGuard, Cloudflare (incl. Zero Trust), External DNS, Route 53, .NET, Python, Node.js

**DevOps Engineer** Feb 2020 – Jul 2021

*IDenTV · United States (remote)*

- Established the DevOps practice from scratch, with CI/CD on CircleCI and TravisCI and Helm charts for ML deployment.
- Containerised ML services for facial identification, aerial detection and object detection, removing environment inconsistencies.
- Engineered GPU sharing across Kubernetes pods (NVIDIA device plugin, MPS), cutting per-inference cost by over 60% for real-time video analytics.
- Packaged a single-node k3s cluster on a GPU-optimised AWS instance with pre-baked AI models as a golden AMI and ISO (Packer, Ansible) for single-command deployment in air-gapped government environments.
- Halved AWS spend from $24K to $12K a month through right-sizing and consolidation onto OpenStack.
- Provisioned EKS, ECS, RDS (Aurora) and ElastiCache on AWS and Oracle Cloud with Terraform, cutting manual setup time 70%.
- Centralised secrets in HashiCorp Vault, AWS Secrets Manager and Oracle Vault, automating credential management.
- Ran Flussonic Media Server for video streaming.
- Ran Milvus for vector storage.
- Load-tested services with Gatling.

**Stack:** AWS, Oracle Cloud, OpenStack, Kubernetes, k3s, Docker, Helm, Terraform, Ansible, Packer, CircleCI, TravisCI, Jenkins, GitHub Actions, HashiCorp Vault, Prometheus, Grafana, ELK, Sentry, InfluxDB, Gatling, MongoDB, PostgreSQL, Redis, Elasticsearch, Kafka, RabbitMQ, Milvus, Flussonic, Contentful, HAProxy, Nginx, OpenVPN, Python, Node.js

**DevOps Consultant (freelance, part-time)** Mar 2025 – Present

*Emergegen AI · Remote*

- Deployed AWS Account Factory for Terraform to vend accounts with security baselines (SCPs) in under 30 minutes.
- Containerised applications for ECS and EKS, cutting deployment times.
- Built ML inference hosting on EKS and ECS with Milvus for vector storage.
- Built KEDA autoscaling driven by RabbitMQ queue depth that scales GPU instances to zero when idle.
- Scaled ingress traffic with the NGINX Ingress Controller.
- Designed cloud-native architectures with the developers, improving reliability.

**Stack:** AWS, Kubernetes (EKS), ECS, KEDA, Terraform, Packer, ArgoCD, Nginx, RabbitMQ, Redis, Milvus, Trivy, Pritunl, Cloudflare, Route 53, GoDaddy, Python

**Blockchain and Data Platform Consultant (freelance)** Freelance

*Twine and other clients · Remote*

- **Twine** (cross-chain Layer 2): deployed high-availability EVM node and validator infrastructure on EKS with sub-second RPC latency and 99.99% uptime.
- **Twine:** automated smart-contract build and deployment pipelines with Hardhat and Foundry.
- **Twine:** added security analysis to smart-contract pipelines before mainnet release.
- Published AWS and Azure apps on the Snowflake Marketplace via Snowflake Container Services.
- Wrote nested CloudFormation templates for metered AWS SaaS Marketplace listings.
- Ran Argo CD for zero-downtime deployments of applications and cluster add-ons.
- Chaos-tested services with LitmusChaos.
- Implemented Cilium eBPF networking.
- Built infrastructure controls for SOC 2.

**Stack:** AWS, Azure, Kubernetes (EKS), ArgoCD, Cilium, LitmusChaos, Snowflake, CloudFormation, Hardhat, Foundry

**DevOps Consultant (freelance, part-time)** Jun 2021 – Oct 2022
*Altada Technology Solutions · Remote*

- Built real-time data pipelines with AWS Glue and Kinesis Firehose, provisioned with Terraform and Terragrunt.
- Designed a multi-account AWS architecture that isolates services by environment for security and compliance.
- Worked with stakeholders in pre-sales to align deliverables with client needs and shape the roadmap.
- Wrote a DevOps strategy from scratch, consolidating tools and planning their integration.
- Built secure development practices into source control, CI/CD and infrastructure management to align with compliance frameworks.

**Stack:** AWS, Kubernetes, Helm, Terraform, Terragrunt, GitLab CI, Prometheus, Grafana, ELK, PostgreSQL, MongoDB Atlas, Redis, Elasticsearch, RabbitMQ, Nginx, Fortigate, CrowdStrike Falcon, Python, Node.js

**Linux System Support Engineer** Apr 2019 – Apr 2020
*NepalLink Network · Kathmandu, Nepal*

- Resolved client hosting issues within SLA: cPanel and ZPanel, email, DNS and SSL certificates.
- Scanned hosted websites for malware, viruses and backdoors with Maldet, ClamAV and CSF.
- Built and configured VPS servers to client requirements, including web servers, databases and monitoring.
- Set up Icinga monitoring, improving response time to issues by 40%.
- Found and fixed performance bottlenecks in system and network resources, improving uptime.
- Documented configurations and fixes, cutting resolution times and onboarding time.

**Stack:** Linux, Bash, LAMP, LEMP, cPanel, ZPanel, WHMCS, Let's Encrypt (Certbot), Icinga, Zabbix, ClamAV, Maldet, rkhunter, chkrootkit, CSF, VMware, Vagrant, OpenVPN

**Junior Network Monitoring Engineer** Apr 2018 – Apr 2019
*Vianet Communications Pvt. Ltd. · Kathmandu, Nepal*

- Analysed system and bandwidth utilisation and optimised traffic.
- Managed network devices and hardware with proactive fault management.
- Ran maintenance with the NOC and fibre teams, minimising service interruptions.
- Resolved escalated network, system and hardware issues.
- Configured switches and OLTs for new POP installations.
- Wrote scripts that automate OLT reboots across regions.
- Kept device documentation for troubleshooting and upgrades, and coordinated incident escalation between teams.

**Stack:** Bash, Icinga, Nagios, Zabbix, Cacti, CheckMK, SmokePing, MRTG

**Technical Researcher** Dec 2016 – Feb 2018
*Datalytics Pvt. Ltd. · Kathmandu, Nepal*

- Delivered client research projects, including voice recognition systems and micro hydro power plants.
- Guided new recruits through project delivery.
- Troubleshot departmental network issues and solved Cisco network assessments.
- Wrote shell scripts that log user access for security monitoring.

**Stack:** Linux, Bash, Cisco networking

## Projects and Open Source

- **keda-gpu-scaler** (Go, Helm, gRPC): collaborator with 13 merged PRs on a KEDA external scaler for Kubernetes GPU workloads; published the Helm chart to GHCR via OCI with chart tests, added coverage reporting and format gating to CI, and extended the `gpu-metrics` CLI.
- **gpu-mcp-server** (Go, MCP, NVML): collaborator with 6 merged PRs on an MCP server exposing NVIDIA GPU metrics to AI agents; built PID-level GPU process attribution, Streamable HTTP transport and a Helm chart for DaemonSet deployment.
- **Autobase** (Ansible, Go, React): PRs to the PostgreSQL HA automation platform, including configurable HAProxy failover so long-running queries survive, Docker Secrets support in the console, and private-subnet deployments across five clouds.
- **Smaller contributions:** screenpipe (fixed the `npx` install of its MCP HTTP server; TypeScript), KaTeX (rewrote the self-hosting documentation) and anchore/scan-action (updated CI workflow examples).
- **AUR**: maintain `tfk8s-bin` (Kubernetes YAML to HCL) and `shellfirm-bin` (risky-command guard).
- **Own projects:** apex (live, byte-accurate visualisations of how computers work, 300+ tested topics), rampart (self-hosted uptime monitoring on Postgres) and terraform-aws-mongodbatlas (Terraform module for MongoDB Atlas).

## Certification and Education

- **AWS Certified DevOps Engineer – Professional**, Amazon Web Services (Credly credential)
- **Bachelor of Engineering, Electrical and Electronics**, Kathmandu University, 2012 – 2016
- **Master of Arts in Psychology** (in progress), Tri-Chandra Multiple Campus, 2023 – Present