

moi@opscraft ~ \$ whoami

Mohamed Ghenaim

Freelance AWS Cloud Architect, DevOps & SRE. Build, run, debug for 9 years.

AWS

Terraform

Kubernetes

AI / LLM

GitOps

CI/CD

Docker

Linux

moi@opscraft ~ \$ fastfetch --logo-type kitty

	focus	Generative AI platforms: air-gapped on-prem LLMs and AWS EKS
	email	mohamed.ghenaim@opscraft.fr
	linkedin	in/ghenaim-mohamed
	location	Bois d'Arcy (78), Île-de-France area, or remote
	languages	French, fluent English (C2, TOEIC 915)
	uptime	9 years in production

[HEAD → mission](#)

June 2026 - present

CA-GIP, Crédit Agricole Group Infrastructure Platform

Cloud & DevOps Architect, sovereign generative AI platform (Mistral AI)

Technical lead for the build of the group's Mistral AI GenAI platform, on-premise on air-gapped OpenShift (NVIDIA B300 GPUs) and on AWS EKS. Driving force behind design and go-live, in a team of four.

@@ [on-prem/openshift-airgapped](#) @@

- + Deployment and version upgrades of the Mistral AI stack in sandbox, DEV and PROD, with a Helm chart adapted to OpenShift.
- + Migration from ArgoCD to FluxCD on my own initiative: per-application GitOps organisation, documentation, team upskilling.
- + Deployment and tuning of models with vLLM (Mistral Small 4, Medium, Large, OCR, Voxtral, Embed).
- + Benchmarks of LLM endpoints with GuideLLM and identification of bottlenecks.
- + Observability of nodes, workloads, vLLM inference and GPUs (Tensor Cores, temperature, power draw) with Grafana and alerting, Dynatrace and Kibana.
- + CNPG PostgreSQL databases backed up to on-prem S3, restores tested and validated.
- + Model and backup storage migrated to on-prem S3, automated model synchronisation in the air-gapped environment.
- + Secrets with HashiCorp Vault and External Secrets, PVC encryption.
- + In-cluster Redis on an image built in GitLab CI, replacing an unsuitable managed service.
- + Go-live of the application and inference clusters: network flows, security validation, platform base, together with the network and security teams and the vendor.
- + Recovery of an internal RAG application after a data-loss incident.

@@ [aws/eks](#) @@

- + Mistral AI stack deployed on EKS across all environments, from DEV to PROD.
- + Exposure through Network Load Balancers, evaluation of EKS Auto Mode.
- + RDS Proxy for connection pooling to RDS databases.
- + Automated creation of Keycloak user accounts.
- + Secrets in AWS Secrets Manager, encrypted with a KMS key backed by CloudHSM.

Stack: openshift, eks, helm, fluxcd, vllm, guidellm, grafana, dynatrace, cnpg, vault, keycloak, kms, cloudhsm

**epex-spot**

Jul. 2025 – Sep. 2025

EPEX SPOT

AWS Cloud Architect, Foundation team

Architecture and deployment of the critical AWS building blocks of a platform being migrated.
Standardised Terraform modules for the DevOps teams.

- + Hybrid Route 53 DNS: central DNS Hub in Terraform, connected to the datacenter through Direct Connect and DX Gateway, shared via Organizations and RAM. Asymmetric routing solved with Network Firewall.
- + Managed PKI with AWS PCA: Root and Subordinate CA hierarchy in Terraform, cross-account sharing, automatic ACM renewal.
- + GitLab Runners infrastructure rebuilt as IaC after a major incident, with no interruption to deployments.
- + GitLab runners module on EC2: autoscaled manager and workers (AWS fleeting plugin), standardised Debian AMI built with Packer.
- + Terraform module catalogue: gitlab-runner, pki-certificate, dns-record, ecr-repository, codeartifact-repository.
- + Tag-based stop and start of resources via Lambda, outside business hours: **-60 %** EC2 cost.
- + Architecture documentation (ADRs), operational guides and AWS upskilling of the team.

Stack: terraform, route53, direct-connect, network-firewall, aws-pca, acm, gitlab-ci, packer, lambda

**cLaranet**

Dec. 2022 – Jan. 2025

Claranet France

AWS CloudOps, Cloud Native BU

Production build and run on multi-account AWS environments for **11** clients: fintech, real estate (CAC 40), streaming, IoT, pharma, banking, e-commerce.

- + L3 support and night on-call across the whole portfolio.
- + Complete ECS stack: PHP Docker images, IAM, ECR, task definitions, RDS, ALB, Route 53, ACM, HashiCorp Vault secrets.
- + Splunk and OpenTelemetry monitoring on EC2 and ECS.
- + Patch management: AMIs migrated from Ubuntu to Debian, GitLab Runners and Docker Machine upgraded.
- + Ansible playbook for application deployment on EC2.

Stack: ecs, ecr, rds, alb, route53, vault, splunk, opentelemetry, ansible



orange

Oct. 2020 - Dec. 2022

Orange France

Linux / DevOps Production Engineer, Portal Common Service

- + Build of the high-traffic URL redirection service: migration from Ubuntu 14.04 to 18.04, Apache, IP addressing, Ansible playbook.
- + L3 support for the Orange Portal production.
- + Shared GitLab CI pipeline templates for all IT teams.
- + Migration of services to Kubernetes and CloudFoundry, Kibana dashboards for production monitoring.
- + DAM application (Nuxeo) migrated to Docker Swarm: optimised Dockerfile, CI/CD pipelines, HTTPS certificates.

Stack: apache, ansible, gitlab-ci, kubernetes, cloudfoundry, docker-swarm, kibana



malakoff-humanis

Sep. 2017 - Sep. 2020

Malakoff Humanis

Linux Infrastructure Production Engineer, CSI/Linux team

- + Linux infrastructure administration: Bash, Nagios, LDAP, Kerberos.
- + Complete ELK cluster (Elasticsearch, Logstash, Kibana, Kafka, Zookeeper) on Docker Swarm.
- + CI/CD automation with Rundeck, Ansible, Docker, GitLab and Jenkins. ITSM with GLPI and FusionInventory.
- + RedHat Satellite (channels, hosts) and Identity Manager.

Stack: rhel, elk, kafka, docker-swarm, rundeck, ansible, jenkins, satellite

Stack

stack.tf

\$ ls ~/.logos



AWS

EKS, Lambda, KMS,
CloudHSM, multi-
account



Terraform

Reusable modules, IaC
for the whole
infrastructure



Kubernetes

OpenShift, EKS, K3s,
GitOps



Docker

Custom images, Swarm,
CI builds



Linux

Debian, Ubuntu, RHEL,
since 2017



Mistral AI

vLLM, sovereign GenAI
platform

```

1  # Skills, by domain
2  resource "skills" "cloud_iac" {
3      aws    = ["EKS", "Secrets Manager", "KMS", "CloudHSM"]
4      tools  = ["Terraform", "Packer", "Ansible", "Hetzner (dedicated)"]
5  }
6
7  resource "skills" "containers" {
8      runtime = ["Kubernetes", "OpenShift", "K3s", "Docker"]
9      gitops  = ["Helm", "FluxCD", "ArgoCD"]
10 }
11
12 resource "skills" "network_security" {
13     mesh    = ["Cilium", "Linkerd"]
14     secrets = ["Vault", "External Secrets", "SOPS", "Infisical", "Keycloak"]
15 }
16
17 resource "skills" "ai_llm" {
18     serving = ["vLLM", "GuideLLM", "LiteLLM", "Mistral AI", "Ollama"]
19     usage   = ["RAG", "embedding pipelines", "n8n AI Agents"]
20 }
21
22 resource "skills" "data" {
23     databases = ["PostgreSQL (CNPg, Barman)", "Redis", "MongoDB", "S3"]
24     vectors   = ["Qdrant", "pgvector"]
25 }
26
27 resource "skills" "observability" {
28     tools = ["Grafana", "Prometheus", "Datadog", "Dynatrace", "Splunk", "New Relic", "ELK"]
29 }
30
31 resource "skills" "ci_cd" {
32     tools = ["GitLab CI", "Artifactory", "Rundeck", "n8n"]
33 }
34
35 resource "skills" "ai_assisted_dev" {
36     claude_code = ["Rust", "Python / FastAPI", "TypeScript / React", "Tauri", "Alpine.js"]
37 }

```

Plan: 8 to add, 0 to change, 0 to destroy.

Personal projects

projects/

README.md: designed, architected, deployed and operated by me. The applications are developed with Claude Code.

Self-hosted GitOps Kubernetes platform

My permanent testbed: a dedicated Hetzner server running K3s, built and operated by hand since 2022, and rebuilt every time my missions taught me a better way of doing things. Everything is described in Git and deployed by ArgoCD, with encrypted secrets and an up-to-date runbook.

30

Helm releases declared in Helmfile

26

applications synced by ArgoCD

3

architecture generations

12 months

of backups: daily, weekly, monthly

v5

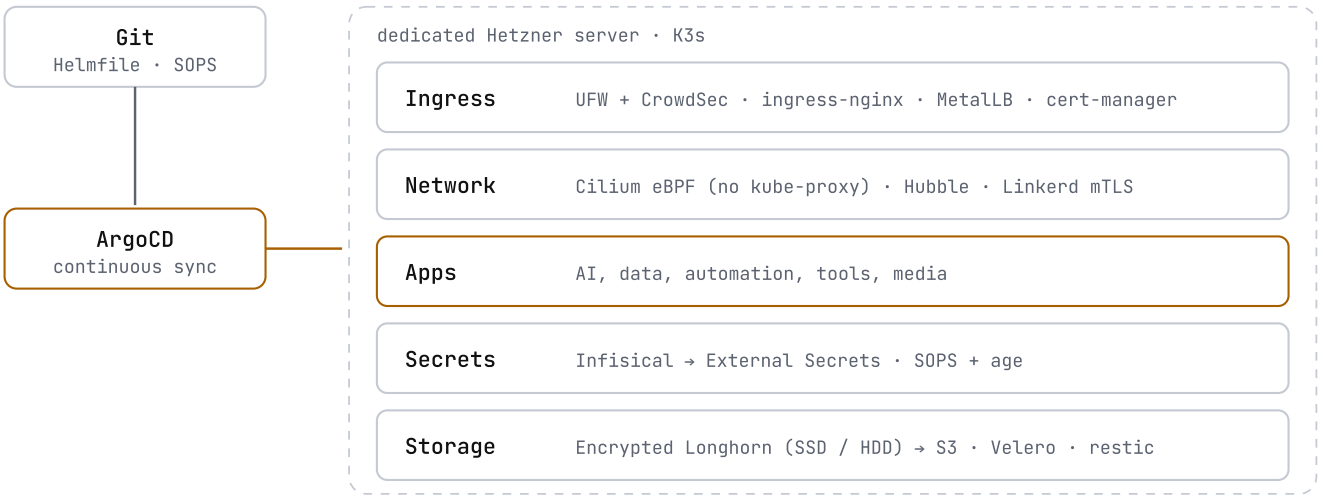
Manifests and Helm
K3s, Traefik, Prometheus and Grafana, first self-hosted services.

v6

GitOps
Helmfile, ArgoCD, SOPS + age secrets, pre-commit hooks, Wiki.js runbook.

v7

All-Terraform experiment
Every app in OpenTofu. Lesson learned: Terraform for infrastructure, GitOps for the app lifecycle.



```
$ argocd app list
GROUP      APPLICATIONS                                SYNC
network    cilium, linkerd, ingress-nginx, metallb     Synced
security    infisical, external-secrets, cert-manager, crowdsec Synced
storage     longhorn, velero                             Synced
data        pgvector, qdrant, mongodb                   Synced
ai          open-webui, ollama, n8n                      Synced
tools       gitea, wiki.js, hedgedoc, affine, matomo     Synced
tuning      vpa, goldilocks                             Synced
media       media-center                                Synced
# excerpt, 26 applications in total
```

Backups and recovery

Longhorn and Velero snapshots to Hetzner Object Storage, files via restic. Daily, weekly and monthly rotation, restore procedures documented in the runbook.

- k3s
- argocd
- helmfile
- sops
- cilium
- linkerd
- longhorn
- velero
- infisical
- crowdsec
- n8n

AI on logs

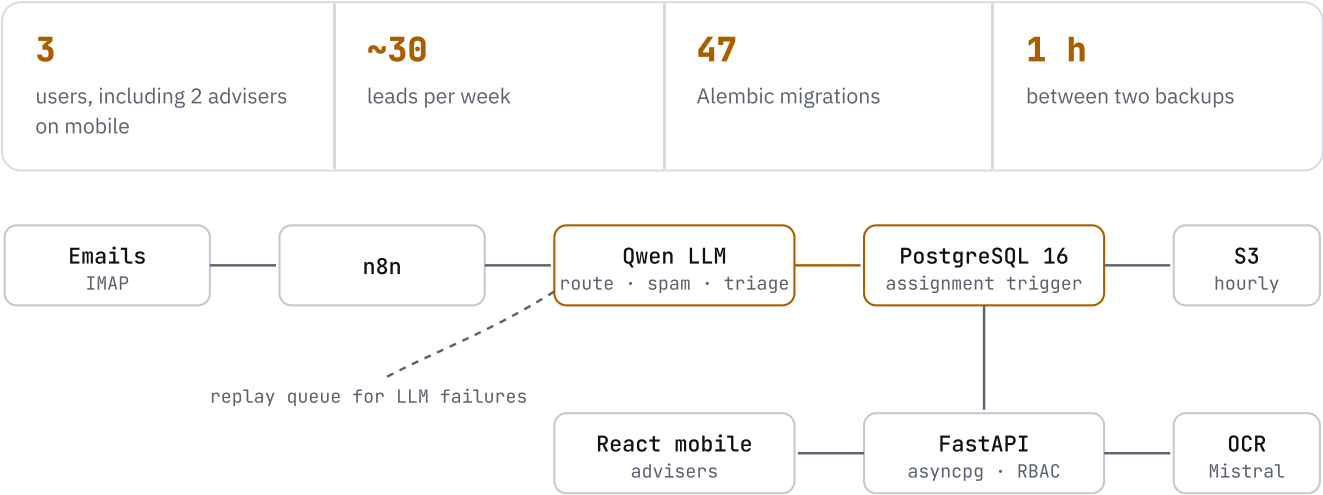
RAG proof of concept on system and ingress logs: logs are embedded, stored in Qdrant and pgvector, then queried through an LLM. Alongside it, an n8n workflow reads the ingress logs, compares them with previous reports and produces a report with an urgency score, committed to Git and flagged on Discord.

projects/crm-immo/

in production since Apr. 2026

Business CRM in production, new-build real estate SME

A web CRM that replaces an Excel tracking sheet. I designed, built and deployed it, and I run it on my own: from the database schema to operations, including the mobile app used by advisers in the field.



Application

FastAPI (asynccpg, Pydantic v2), PostgreSQL 16, React 18 and TypeScript (TanStack Query, Tailwind), designed mobile first. Role-based RBAC, Excel and PDF exports.

Deployment

Production on docker-compose on a Scaleway Dedibox, pre-production on K3s (CNPG, Longhorn). Rootless Buildah images, healthchecks, smoke tests, rollback procedure.

- fastapi
- postgresql
- alembic
- react
- typescript
- n8n
- llm
- buildah
- pytest

AI and automation

Leads ingested and classified by an LLM, OCR of supporting documents, weighted assignment to advisers directly in the database.

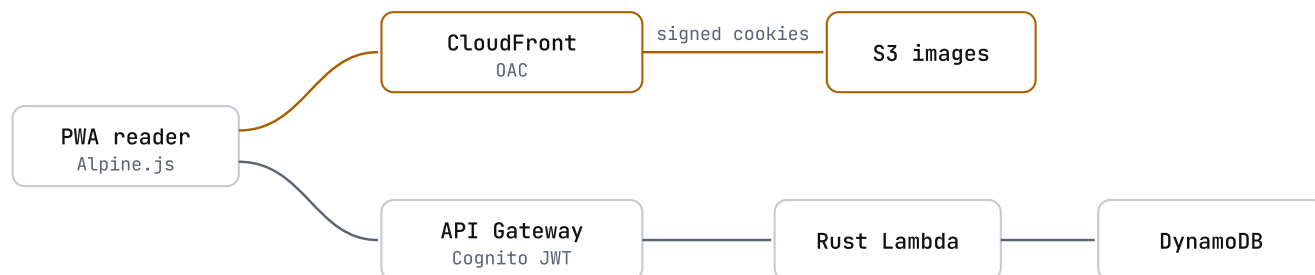
Resilience and quality

Hourly backups to S3 with a Discord alert on failure. pytest tests against an ephemeral PostgreSQL database in K3s, vitest on the front end.

Manga Shelf, a serverless manga reader

A reader in production, 100% serverless on the AWS Free Tier, with the whole infrastructure in Terraform. While reading, images go through CloudFront with signed cookies: no Lambda is invoked.

34 API Gateway routes	7 DynamoDB tables	0 Lambda invocations while reading	Free Tier 100% serverless
---------------------------------	-----------------------------	--	-------------------------------------



Site and API

Rust Lambda on arm64, Cognito PKCE with admin groups, CloudFront private key in SSM. Alpine.js PWA front end, Puppeteer UI tests and k6 load tests.

Admin application

Rust desktop application (WebView2): publishing to S3 and CloudFront invalidation, built-in BitTorrent client, userspace WireGuard tunnel, chapter titles extracted by OCR through the Claude API.

rust lambda arm64 api-gateway dynamodb cognito cloudfront terraform k6

Secret Vault, one-time secrets

Zero-knowledge encryption in the browser (AES-256-GCM, PBKDF2 with 200,000 iterations). The key stays in the URL fragment and never reaches the server. Three Rust Lambda functions, atomic destruction on read, expiry through DynamoDB TTL, API Gateway throttling, Playwright E2E tests.



rust webcrypto dynamodb ttl playwright

projects/pulse-guard/

in production

Pulse Guard

URL monitoring every minute through EventBridge and a Rust Lambda: HTTP status, latency, TLS expiry. Discord alerts, public status page, native Cognito JWT authorizer on API Gateway.

rust

eventbridge

cognito jwt

discord

projects/desktop-tools/

since Mar. 2026

Windows desktop tools in Rust

Chuchotement: push-to-talk voice dictation, 100% local recognition (Parakeet).

Note de frais: Tauri 2, receipt OCR through the Claude API, PDF and Excel reports.

Texto: streaming proofreader, word-level diff, cost tracked per request.

Recommendations

reviews.md

✓ approved

Florent Michel Head of SysOps / Product Owner, EPEX SPOT

« Mohamed est intervenu chez nous pour un support ponctuel sur AWS afin de nous aider à respecter des deadlines critiques. Malgré la durée limitée de sa mission, il a su démontrer un haut niveau de compétence et d'efficacité. Son expertise, sa réactivité et son implication ont été des atouts indéniables pour la réussite de nos projets. »

✓ approved

Grégory Daudin Product Thinking & Agile Delivery, EPEX SPOT

“He played a key role in helping us catch up on significant AWS technical debt and even saved us from a really complex situation that could have blocked critical progress. We would have loved to keep him longer in the team. It was a pleasure!”

Education

education.txt

2018 – 2020

Master of Science (M2) in Cybersecurity ITESCIA, Cergy

2017 – 2018

Bachelor's degree (Licence) in General Computer Science CNAM

2015 – 2017

BTS Services Informatiques aux Organisations (IT services for organisations) Plaisir

```
$ ./contact.sh
```

A platform to build, or to get back on track?

AWS architecture, Kubernetes, generative AI: describe your context by email.

served by S3 and CloudFront, deployed with terraform apply

JetBrains Mono and IBM Plex Sans fonts under the SIL OFL 1.1 license. Logos from the dashboard-icons repository (Apache-2.0 license); trademarks belong to their respective owners.