Guido Maria **Serra**

Forward Deployed Engineer · AI Infrastructure & Sovereign Deployments

+49 173 7349 795 | guido@whitehat.berlin | Berlin, Germany | github.com/zeph | linkedin.com/in/zeph1ro

Work Authorization: *EU Citizen — eligible for US & UAE work visa sponsorship.*

Summary

Engineer who deploys AI where it can't touch the cloud: air-gapped LLM platforms, GPU Kubernetes, regulated and sovereign environments. 30-year trajectory: a decade of direct client delivery (UNESCO, banks, telcos), then Staff-level SRE and AI platform engineering, with fractional technical leadership along the way. Known as a capability enabler: hired and built teams from scratch, coached SRE organizations, guided companies through restructurings — while staying hands-on. Deep security and identity practice (SIEM, IAM, Kerberos/OIDC) across bare metal, KVM, OpenStack and Kubernetes. After eight years of referral-driven fractional work, now looking to go deep with one product and one team.

Selected Impact

- Operated PlayStation Now infrastructure: **55,000 hosts across 12 datacenters**, monitored fleet-wide via Graphite + SNMP.
- Built observability where none existed: Prometheus/Grafana/Datadog platforms, business-level telemetry (Sony, Mobivia, HelloFresh).
- **100% referral-sourced client base** since 2018 — every engagement via long-standing client and colleague relationships.
- **$300K/year recurring savings** delivered on a $30M+/year infrastructure spend portfolio (OLX FinOps).
- **16 consecutive contract renewals** over a 4-year engagement (Mobivia).

Core Competencies

Observability: Prometheus, Grafana, ELK, Datadog, Graylog, OpenTelemetry
Cloud Native & AI: Kubernetes (GPU), ArgoCD, Traefik, SealedSecrets, Cilium, Terraform, ClickHouse/MinIO, air-gapped LLMs
Networking, Edge & CDN: F5 BIG-IP (LTM/iRules), Cloudflare, Akamai, Fastly, DNS architecture, load balancing, WAF
Security & Identity: Keycloak, OIDC, LDAP/Kerberos, Active Directory, Vault, SIEM, DMZ design, HLD/LLD documentation
On-Prem & Storage: KVM (Ganeti), OpenStack, VMware, Ceph, MinIO, bare-metal provisioning (PXE), datacenter operations
Leadership & Budget: Team hiring & building, coaching, handover, stakeholder management, $30M+/yr vendor & budget mgmt

Open Source & Community

Open Source: FSFE Fellow; Debian maintainer; Traefik contributor; founder, Politecnico Open *nix Labs (POuL).
Security Community: SOC GOON at DEFCON 2025; iCTF World Champion 2005.

Professional Experience

Contractor — whitehat.berlin | fractional and embedded engagements | Berlin / Remote (2018 – Present)

- **Current partnerships & engagements** *(2022 – present)* — all referral-sourced from long-standing client and colleague relationships: DFND Security (EMEA Solutions Partner: Fastly Signal Sciences ngWAF architecture); Forward Earth (SOC 2, ISO 27001 and Vanta); Contensi Software (DevOps partnership: MinIO/Rancher, CDKTF); 3T Software Labs (Kotlin DSL CI/CD engineering; re-engaged in 2026).
- **Stealth FinTech Startup** *(Dubai, UAE, Remote)* – Senior Software Development Engineer, *Feb 2025 – Oct 2025*: tokenized financial infrastructure in Python and Node.js/TypeScript; FastAPI microservices; AI-native engineering workflows (Cursor, Windsurf).
- **Octostar** – Head of DevOps / SRE, *May 2024 – Feb 2025 (Bergamo, Hybrid)*:
  - Led engineering team deploying multi-architecture GPU Kubernetes clusters for LLM training/inference on embedded/portable devices; ClickHouse/MinIO data platform.
  - Prepared infrastructure for fully air-gapped environments with offline artifact handling; delivered a local-LLM solution.
- **Max Planck Institute** – Site Reliability and Security Engineer, *Dec 2022 – May 2024 (Berlin, Hybrid)*: hybrid OpenStack/Kubernetes platform with **1.5PB MinIO** storage; federated IAM (LDAP/SAML/OAuth2); IaC with CDKTF; helped the organization hire an FTE.
- **Prima Assicurazioni** – Staff Site Reliability Engineer (Fractional), *Nov 2022 – Jul 2023 (Milan, Remote)*: guided a 12-engineer SRE team on blameless culture and incident response; coordinated AWS multi-region rearchitecture (London, Barcelona, Milan); infra scaling, cost optimization and governance; stakeholder management through the DDD restructuring.
- *2023* — architecture due-diligence audit for investors (Bella & Bona); embedded Debian/armhf development for IoT charging infrastructure (SkyCharge).
- **ClovrLabs** – Cloud Architect, *Sep 2021 – Jul 2023 (Barcelona, Remote)*, blockchain reputation unicorn: source-of-funds risk across the entire public BTC transaction graph (5-node Redis); AWS EKS with Cilium network policy; built a 5-person security ops team; federated IAM (SAML/Google/Cloudflare Access).
- **HelloFresh** – Security / DevOps Engineer, *Apr 2020 – Mar 2021 (Berlin, Hybrid)*: 2-week emergency extended into a multi-year engagement; stabilized the SIEM platform (ELK + S3 ingestion) with retention policies controlling cost/sizing; library vulnerability scanning with Jira integration; hired and onboarded a full FTE team.
- **Volkswagen WeShare** – DevOps Team Lead, *Mar – Dec 2019 (Berlin)*: brought in one week before go-live to ensure production readiness; led the 5-person SRE team through launch; microservices on GCP/Azure.

- **Mobivia (ATU/Norauto)** – DevOps Engineer, *Oct 2018 – Dec 2022 (Berlin, Remote)*; 4-year engagement, rolling 3-month contracts
  - Unwound an organically grown on-premise Kubernetes setup; led Link11 → Cloudflare edge migration (DNS restructuring, DDoS/bot protection via DataDome, **DMZ hardening**); HLD/LLD documentation, test/cutover/rollback plan; CDN/DNS/network SME.
  - Migrated three script-managed NGINX VMs to Traefik after isolating all traffic flows — live rule dashboard for devs.
  - GitOps (ArgoCD + SealedSecrets) over undocumented Helm installs and Ansible-encrypted secrets; granular Kubernetes RBAC.
  - Logging & IAM: redesigned **Logstash** rulesets, tuned **ELK** retention; per-booking business telemetry via **Datadog**; Keycloak (OIDC) + AD (LDAP/Kerberos); SPNEGO via in-cluster NGINX behind Traefik.
- **SiWeGO** – Fractional CTO, *Nov 2017 – May 2019 (Remote, alongside FTE roles)*: technical leadership for a B2B logistics startup; designed the cost-optimized architecture; co-authored a patent application for algorithmic graph indexing.

Full-Time Employment (2010 – 2018)

OLX Group

Senior Performance Engineer — *Sep 2016 – Jun 2018*

- Web performance domain expert; load & performance testing across 4 locations (AR, PT, DE, IN); CDN/SSL/DNS vendor selection and coaching.
- Managed **NewRelic ($2M)**, **Akamai ($4M)**, and **AWS ($24M)** contracts; FinOps practices across the portfolio.
- Contributions to the Goad AWS Lambda load-testing tool (Go).

Sony PlayStation (Gaikai → PlayStation Now)

Site Reliability Engineer — *Oct 2013 – Jun 2016*

- Operated the global game-streaming fleet for PlayStation Now; on-call across the Atlantic; mentored 5+ engineers.
- Drove org-wide **Grafana** adoption by building a textX DSL on graphite-api — metrics self-service for engineering teams; OS-level performance tuning for the storage tier.
- Designed NTP infrastructure for low-latency gaming; PKI-based SSH platform; architected Ceph storage (RBD, CephFS, S3/RGW); PXE-booted Tokyo and London DCs; resolved DHCP/VPN MTU issues.

Rocket Internet

Chief Troubleshooter / Senior Data Architect — *Oct 2012 – Dec 2013*

- Direct report to the Global CTO; resolved production issues across 300+ ventures globally; built a Hadoop/MapR cluster for big-data analytics.
- Author of SqlHBase (pypi.org/project/SqlHBase); patched a major PHP data structure (adjacency tree in pure SQL with stored procedures).
- Solved TCP window scaling on high-latency networks (stunnel + rsyncd).

txtr GmbH (3M-sponsored)

Backend QA Manager — *Jul 2011 – Aug 2012*

- Owned site reliability for Europe's largest content aggregator; centralized logging via **Graylog**; JMX-instrumented metrics; Ruby/BDD (Cucumber) test automation.

Nokia Maps (HERE) & ProfitBricks (IONOS)

Senior System Engineer / Administrator — *2010 – 2011*

- Operations and R&D for OVI Maps; CI environments (Puppet, CentOS); KVM-based IaaS cloud operations.

Contract (2008 – 2010)

Vodafone Group — *Düsseldorf, Germany*

QA Architect — *Jan 2008 – Jun 2010*

- **10 consecutive contract renewals** over 2.5 years (rolling 3-month contracts) during the Italy → Germany transition.
- Led load & performance testing; identified major architectural flaws in Java enterprise applications; authored an internal load-testing toolkit (HTTP-QAT) and stress-tested its 360 backends; remote GSM/3G monitoring checks across networks and SIMs.

Independent Engineering – Italy (1996 – 2008)

Solution Architect and Systems Engineer | Multiple Clients. From fixing dial-up connections and building client websites (1996) through Flash/ActionScript development to systems engineering from early 2001 — scoped, budgeted, delivered, and coached the handover, before the industry had a name for it.

- **UNESCO** (UN agency, two engagements): geospatial DB query-wizard delivered in **5 working days** (client's prior in-house attempt: 3 months); MIPS-based remote-controlled gate system on embedded Linux.
- **Telco, banking & energy:** security analysis at H3G (Cisco PIX, CyberARK, Nessus); Solaris → Linux migration of banking-trading procedures (BPU Banca); zero-documentation HA billing-cluster OS upgrade — userspace backup and cutover in **7 days** (Sorgenia); undocumented firewall recovery (IANOMI); **3-day turnaround** (Wind).
- **Delivery with handover:** LDAP-based SSO network/VoIP infrastructure for banks and training companies — budgeted, built, and coached client admins to operate (INAS-CISL, Valore Formazione, OmniGATE); software suites for tax-advisor firms (Solidar Servizi); p2p VPN video-broadcast network for a pharma advertiser (led 3 technicians, VideoMAG).
- **Embedded & R&D:** Linux embedded R&D with Alessandro Rubini (*Linux Device Drivers*) and Marco Cavallini (KOAN); Qt/SQLite on ARM for automotive weighing systems (LabEL); C++ protocol redesign for Bticino home automation.
- **Software:** Pandora (SMB stack, heartbeat/DRBD, OpenWRT); Tempus (SMS engine); Stendhal (full-text, digitami.it/stendhal).
- Lecturer in C socket programming, Politecnico di Milano; volunteering: Ingegneria Senza Frontiere.

Certifications & Languages

Certifications: Red Hat Ceph Storage (CEPH125), 2016; ISO 27001 / Vanta Auditor, 2025.
Languages: Italian (Native), English (Fluent), German (Fluent, spoken). **Interests:** Offshore sailing (Centro Velico Caprera, Les Glénans), skiing, cooking, trekking; Alpine Guide (CAI Bergamo).

+39 348 4313 992 | zeph@guidoserra.it · zeph@fsfe.org | Predore, Italy

Autorizzo il trattamento dei miei dati personali presenti nel CV ai sensi del Decreto Legislativo 30 giugno 2003, n. 196 e del GDPR (Regolamento UE 2016/679).