

Kirill Makhonin

Principal Software Engineer / Platform Architect

Distributed systems & ML/data infrastructure

New York, NY | career@makhonin.biz | [LinkedIn](#) | [cv.makhonin.biz](#)

Hands-on software engineer and platform architect **building distributed data, ML/AI, and infrastructure platforms** across heterogeneous environments. **Use agentic development workflows**, primarily Claude Code, across architecture and codebase exploration, implementation, testing, and review, while retaining ownership of technical decisions and production quality. **Build production systems with Python, Go, TypeScript/React**, and a broad range of IaaS tooling.

17+ years of experience spanning industrial simulation and geospatial applications, enterprise data platforms, distributed ML training and data platforms, identity systems, GPU infrastructure, and AI application platforms.

EXPERIENCE

Platform Architect

US Digital Infrastructure Company | Feb 2026 - Present

- Architected and built GPU compute infrastructure supporting virtual machines and physical servers. Owned system design and hands-on implementation across resource provisioning, management, and remote access.
- Lead technical delivery of an AI application platform with engineering and product. Built isolated execution environments and service integrations while overseeing the agent core and data subsystem.
- Designed and led implementation of security architecture for a multi-tenant platform. Covered identity, authorization, workload isolation, and data protection across services and infrastructure.
- Integrated an open-source inference engine and designed and built request routing and load balancing for language-model workloads.
- Built an initial prototype for distributed firewall and overlay-network functions using hardware acceleration. Connected platform software with programmable networking hardware.

Sr. Principal Engineer, Platform Architect

US Agricultural Technology Company | Nov 2022 - Jan 2026

- Owned technical direction, development, and operations for a platform infrastructure and reliability team supporting data and ML engineers. Established on-call, service objectives, monitoring, and support practices.
- Re-architected shared analytics software into an autoscaling, multi-tenant platform. Migrated large data collections while preserving compatibility and service continuity for existing workflows.
- Built distributed computer-vision training workflows across cloud and on-premises environments. Integrated different training systems, networking, and workload placement with shared orchestration and developer tooling.
- Developed a research prototype into a hardware-in-the-loop testing platform. Automated equipment allocation, recovery, monitoring, and simulated inputs for application and system-software validation.

Infrastructure Tech Lead

US EV Manufacturer (through EPAM Systems) | Jun 2021 - Nov 2022

- Led architecture and engineering across a joint client and EPAM team. Built data-ingestion and model-workflow blueprints, implemented selected workflows, and integrated data-quality monitoring.
- Created a shared infrastructure SDK and CLI for reusable provisioning and deployment components. Improved pipeline development with common abstractions, workflow tooling, and validation before deployment.

EXPERIENCE / CONTINUED

Infrastructure Tech Lead

US Credit Bureau (through EPAM Systems) | Aug 2019 - Jun 2021

- Led engineering for internal and customer-facing analytics platforms from implementation through maintenance. Helped build the engineering team and owned technology decisions with product.
- Owned architecture and implementation of a cloud analytics platform, including redundant infrastructure, interactive development environments, automated provisioning, and service recovery.
- Led platform security design and data-protection work. Supported internal certification, independent security audits, and compliance documentation.

Lead Systems Engineer

EPAM Systems | Sep 2017 - Aug 2019

- Led implementation of a Healthcare ML platform through production delivery. Coordinated engineers, an architect, and the customer on data-processing and model workflows.
- Built and released open-source ML training and deployment software. Connected notebook-based development with training infrastructure, model export, and deployment workflows.
- Helped establish Python engineering within the data-engineering practice. Developed technical training for external participants, including curriculum, assessments, and mentoring; graduates joined the company.

Senior C++ Developer, Architect

Russian University Research Laboratory | Jun 2012 - Aug 2017

- Owned architecture and end-to-end development of an offline Arctic monitoring application. Integrated radar, vessel positions, and satellite imagery with data processing, databases, and visualization for remote field use.
- Built networking, communication, and deployment layers for industrial training simulators. Modernized a C++ environmental-simulation core, added capabilities, and automated testing.

Web Developer

Crocus City Hall | Sep 2009 - Jun 2012

- Delivered and supported a full-stack online ticketing and reservation platform for a concert venue. Owned the customer interface, backend, and booking workflows, starting the commercial project while in high school.

TECHNICAL EXPERTISE

Languages: Python, Go, TypeScript/React, C/C++, C#/ .NET, Rust

Infrastructure: Kubernetes, Kubernetes webhooks & operators / Operator SDK, AWS, GCP, Terraform, Terraform providers, Terragrunt, Pulumi, Ansible, Cloud Hypervisor, NetBox, HA, Load Balancing, Linux, MacOS, HELM, Nvidia GPU Operator, PXE/iPXE, Multus, IPMI, Redfish, Modbus

ML and data: Kubeflow, Slurm, EFA/RDMA, Airflow, Databricks, JupyterHub/Lab, Monte Carlo, Kafka, Redis, PostgreSQL, Graph Data Bases, vLLM, MLflow, Claude Code, MCP (development and consuming)

Identity and AI: OIDC, mTLS/PKI, JWT/JWKS, OPA, Firecracker, MCP

Operations: Observability, Prometheus, Grafana, VictoriaMetrics, OpenTelemetry, Loki, PagerDuty, Rootly, GitHub Actions, Jenkins, Agentic Development, Agentic Based Triaging and Operations

Networking: DHCP, DNS, Rest, ConnectRPC, gRPC, MetalLB, Traefik, Istio/Envoy

Security: PKI / HSM, JWT, OAuth/OIDC, Open Policy Agent, OpenBAO, TLS/mTLS

EDUCATION & PROFESSIONAL DEVELOPMENT

ITMO University: M.S. Computer Science with honors, 2018; B.S. Computer Science, 2016.

AWS Certified Machine Learning - Specialty, earned 2021. Architecting with Google Cloud Platform Specialization, 2019.

Languages: English (fluent), Russian (native). Interests: aviation (private pilot), IoT, 3D printing.