

MUHAMMAD FARZAN ALI KHAN

Senior Software Engineer — Backend, Distributed Systems & Applied AI

+92 316 148 0849 | farzankhan009@gmail.com | linkedin.com/in/farzankhan09 | github.com/Farzankhan009 | Islamabad, Pakistan — Open to Remote & Relocation

PROFESSIONAL SUMMARY

Senior Software Engineer with approximately 5 years of experience (October 2021 to present) architecting distributed backend systems, event-driven data pipelines, and applied AI infrastructure across ad-tech and SaaS domains. At Advergić — Pakistan's only Google-Certified MCM Channel Partner — builds and operates a 10+ service microservices platform spanning Go and TypeScript, including fraud/invalid-traffic detection, real-time content classification, publisher and advertiser management systems, and a full self-hosted observability stack (Prometheus, Grafana, Loki, Tempo). Personally trains and deploys production ML models (CatBoost, LightGBM) for real-time yield and floor-price prediction, and has built a production RAG pipeline (LangChain, LangGraph, vector databases) for AI-assisted product features. Concurrently contributes to backend delivery at Code Freaks, where a source-code audit documents 640+ production API endpoints across 10+ client platforms (311 in a single flagship recruitment SaaS product). Deep expertise in NestJS, Go, Kafka, ClickHouse, PostgreSQL, Redis, and vector/RAG systems. Open to senior backend, distributed systems, or backend/AI infrastructure roles globally.

PROFESSIONAL EXPERIENCE

Senior Software Engineer

Advergić — Pakistan's only Google-Certified MCM Channel Partner (Ad Tech, SSP & DSP)

January 2026 – Present | Islamabad, Pakistan (On-site, Full-time)

Tech Stack: NestJS, Go (Golang), Bun.js, Python, Java, Gin, Hono, Apache Kafka, Redpanda, Apache Flink, ClickHouse, PostgreSQL, MongoDB, Aerospike, Qdrant, MinIO, Redis (Streams, RedisSearch, RedisJSON), BullMQ, Prometheus, Grafana, Loki, Tempo, cAdvisor, Caddy, CatBoost, LightGBM, Google Gemini API, OpenAI API, Google Ad Manager API, Google Analytics 4 API, MaxMind GeoIP2, Docker, GitHub Actions

- Architect and maintain a 10+ service distributed microservices platform (within a broader system of roughly 15 to 20 interlinked production services) powering a programmatic ad-monetization exchange serving 200+ publishers and processing hundreds of thousands of ad events daily.
- Contributed to a Go-based invalid-traffic and fraud detection engine (a rewrite of an earlier Node.js system) performing multi-layered bot detection via browser fingerprinting, IP reputation scoring, and click-timing biometric analysis; uses MongoDB for fraud-state documents, Redis Streams with consumer groups to decouple instant scoring from async database writes, Kafka for durable telemetry streaming to downstream ML pipelines, and Aerospike for sub-millisecond fraud-state reads.
- Built a Go-based Google Ad Manager yield-data harvester authenticating via OAuth2 service-account JWT, ingesting hourly programmatic yield reports into ClickHouse (using ReplacingMergeTree for deduplication) across device, country, and ad-unit dimensions.
- Trained and deployed CatBoost and LightGBM models (full pipeline: data preparation, feature engineering, labeling, train/evaluate/test) predicting eCPM and floor-price recommendations from publisher, DSP, GAM, and GA4 data; served via a Redis-cached, sub-millisecond endpoint consumed by a browser-side script for real-time programmatic floor-price suggestions during live bidding.
- Built a real-time content-intelligence service (Bun.js/TypeScript) that scrapes and classifies publisher page content against IAB taxonomy using Google Gemini, with Redis Stack (RedisJSON + Streams) handling cache-miss events without blocking the ad auction, Qdrant vector search enabling semantic similarity matching, and ClickHouse providing a deduplication layer to control LLM API costs.
- Contributed to a multi-service NestJS publisher-management platform (18+ modules, dual-PostgreSQL architecture) with BullMQ FlowProducer-driven onboarding workflows, Socket.io real-time status updates, and dedicated workers ingesting Google Analytics 4 and Google Ad Manager reporting data into a React 18 / Vite / Zustand / TanStack Query analytics dashboard.
- Contributed to an advertiser campaign-management platform using event-driven Kafka messaging to decouple campaign submission from an OpenAI-powered IAB taxonomy classification microservice (Puppeteer-scraped landing pages), with Stripe-based campaign billing and MinIO object storage for ad-creative assets.
- Built and maintained a self-hosted observability stack (Grafana, Prometheus, Loki, Tempo, cAdvisor, Grafana OnCall) with custom dashboards for bid rates, fraud metrics, publisher revenue, and service health, plus Slack/WhatsApp incident alerting.
- Leads code reviews, pull request governance, and deployment pipelines for a team of 5 to 6 backend engineers, setting architecture and performance standards across all active services.

Fullstack JavaScript Developer

Code Freaks — Digital Product Agency (UI/UX, MVP, QA, Tech Consulting)

January 2022 – Present | Remote

Full-time, Jan 2022 – Jan 2026; transitioned to part-time, remote from Jan 2026 to present

Tech Stack: NestJS, Express.js, MySQL, MongoDB, TypeORM, Sequelize.js, MikroORM, LangChain, LangGraph, ChromaDB, React.js, Redux Toolkit, Redux Saga, React Native, Vue.js 2, Vuex, Vuetify, Stripe, Stripe Connect, Twilio, Firebase, AWS S3, Docker, GitHub Actions CI/CD, Passport.js, JWT, Puppeteer, BullMQ

- Led backend development across 10+ client project lifecycles over 4+ years; a source-code audit documents 640+ production REST API endpoints across these platforms, including 311 endpoints in a single flagship recruitment SaaS product, built in NestJS and TypeORM with ACL-based access control and subscription-tier enforcement.
- Built a production Retrieval-Augmented Generation (RAG) pipeline for an AI assessment advisor: LangChain and LangGraph for stateful multi-turn conversation management, ChromaDB as the vector store, and Google Gemini for response generation — enabling recruiters to semantically search hundreds of assessment modules conversationally instead of manual filtering.
- Engineered three dedicated BullMQ job queues for AI proctoring analysis, AI-driven cultural-fit scoring, and post-assessment report generation, plus server-side Puppeteer PDF generation for branded, printable candidate scorecards.
- Resolved a production concurrency bug (duplicate assessment submissions from rapid double-clicks or network retries) by engineering a Redis-based distributed lock (SET NX EX pattern) guaranteeing single-write consistency.
- Integrated Stripe subscription billing (PaymentIntents and tiered Subscriptions with webhook-driven access control) across 6+ products, and Stripe Connect multi-party payouts (OAuth Express onboarding, automated split payments, KYC) for a mentorship marketplace platform.
- Built a mentorship-marketplace backend integrating Google Calendar API for session scheduling, Twilio for WebRTC video sessions and SMS reminders, Google OAuth for one-click signup, and Firebase Cloud Messaging for mobile push notifications — plus a companion React Native mobile app with native Twilio video and Apple/Google Pay support.
- Delivered a peer-to-peer EV-charger booking marketplace with multi-strategy Passport.js authentication (Facebook, Google, JWT) and geo-coordinate-based proximity discovery, and a cryptocurrency community platform with tiered Stripe subscriptions and Google Calendar-integrated session booking.
- Built a cloud print broker using the Google Drive API as a zero-cost document-to-PDF conversion engine and the IPP protocol for direct, driverless communication with network printers, with ephemeral file handling (UUID-renamed, auto-deleted post-conversion) for user document privacy.
- Drove production-grade performance optimization (query indexing, database partitioning, Redis caching), improving response times by up to 40% on critical endpoints; managed and mentored up to 5 backend developers across concurrent project tracks.

Backend Intern (Node.js)

RB Sol — Technology Solutions Company

October 2021 – December 2021 | Lahore, Pakistan

Tech Stack: Node.js, Express.js, MongoDB, Mongoose, Sequelize

- Developed backend CRUD services for small-scale applications using Express.js, managing database interactions through Mongoose and Sequelize ORM.
- Assisted in debugging production codebases and resolving backend issues, developing foundational Node.js skills in a professional environment.

TECHNICAL SKILLS

Languages & Runtimes: TypeScript, JavaScript, Go (Golang), Python, Java, Node.js, Bun.js, SQL

Backend Frameworks: NestJS, Express.js, Gin (Go), Hono (Bun), RESTful API design (640+ endpoints audited in production), Microservices Architecture

Streaming & Event Infrastructure: Apache Kafka, Redpanda, Apache Flink, Redis Streams (Consumer Groups), RedisSearch, RedisJSON, BullMQ (incl. FlowProducer DAG workflows), WebSockets, Socket.io

Databases & Storage: ClickHouse (materialized views, ReplacingMergeTree), PostgreSQL, MySQL, MongoDB, Aerospike, Redis, Qdrant (vector DB), ChromaDB (vector DB), MinIO (object storage), TypeORM, Sequelize.js, Mongoose, MikroORM

AI / ML & Applied Data Science: CatBoost, LightGBM (trained production tabular models), LangChain, LangGraph (RAG pipelines), BGE-M3 (multilingual embeddings), CLIP/mCLIP (image classification), Gemma (text classification), Vector search & embeddings, Prompt engineering, Google Gemini API, OpenAI API, OpenRouter, Modal (serverless GPU), Puppeteer (scraping pipelines)

Ad Tech: Google Ad Manager API (GAM), Google Analytics 4 API, SSP & DSP concepts, IAB content taxonomy classification, IAB TCF 2.x GDPR consent validation, publisher KPI analytics, eCPM/floor-price modeling, MaxMind GeoIP2

Cloud, DevOps & Observability: Docker, Docker Compose (multi-service stacks, 15-30+ services), Alpine Linux, Caddy (reverse proxy), AWS EC2/S3, Firebase (Hosting, Realtime DB, FCM), GitHub Actions CI/CD, Prometheus, Grafana, Loki, Grafana Tempo, cAdvisor, Grafana OnCall, Flyway

APIs & Third-party Integrations: Stripe (Payments, Subscriptions, Connect multi-party payouts), Twilio (WebRTC Video, SMS), Firebase Cloud Messaging, Mailjet, MJML, Cloudinary, Google Drive API, Google Calendar API, Google OAuth, Internet Printing Protocol (IPP)

Auth & Security: OAuth 2.0, JWT (access/refresh rotation), Passport.js (multi-strategy: Google, Facebook, JWT, Local), Sessions, ACL/RBAC (custom decorators), ephemeral file security

Frontend: React.js, Redux Toolkit, Redux Saga, Zustand, TanStack Query, Vite, Vue.js 2, Vuex, Vuetify, React Native, Chakra UI, Material UI, Bootstrap, ApexCharts/Chart.js/Recharts

Testing & Tooling: Jest, Supertest, Zod, ESLint, Prettier, Git/GitHub, Pino/Winston structured logging

KEY PROJECTS

Dexta — AI-Powered Recruiter Assessment Platform (Aug 2023 – Present)

NestJS, TypeORM, MySQL, LangChain, LangGraph, ChromaDB, Google Gemini, BullMQ, Puppeteer, Redis, Stripe, React.js, Redux Toolkit

Comprehensive SaaS recruitment platform automating candidate evaluation via AI-generated tests, proctored sessions, and a conversational AI assessment advisor. 311 audited backend endpoints across 50+ TypeORM entities. Built the RAG-based advisor (LangChain/LangGraph + ChromaDB + Gemini), three BullMQ processing queues (proctoring, cultural-fit AI, report generation), server-side PDF scorecard generation, and a Redis distributed lock resolving a submission race condition.

CareerPaths — Mentorship Marketplace (Web, Mobile & Dashboard)

Express.js, Sequelize, MySQL, Stripe Connect, Twilio, Google Calendar API, Google OAuth, Firebase FCM, React Native, Vue.js

Dual-sided marketplace connecting mentors and mentees for coaching and mock interviews, served via web, Vue.js dashboard, and React Native mobile apps. Built Stripe Connect multi-party payouts with OAuth onboarding, Google Calendar-integrated session booking, Twilio WebRTC video and SMS reminders, and Firebase push notifications.

Winchit — Vehicle Relocation App (Jan 2021 – Present)

Express.js, MongoDB, Mongoose, Stripe, Firebase FCM, Socket.io, Twilio

Real-time vehicle dispatch and relocation platform with instant and scheduled bookings. Led backend development including Stripe setup-intent payments, Twilio messaging, and live WebSocket-based driver tracking.

Dispatcher — Fleet Management Extension (Dec 2023 – Present)

NestJS, MongoDB, Swagger API, AWS S3

Standalone dispatcher service sharing a live MongoDB instance with Winchit. Solved cross-environment database synchronization using custom NestJS RBAC decorators.

EVCSHare — EV Charger P2P Sharing Marketplace

Express.js, Sequelize.js, MySQL, Passport.js (Facebook/Google/JWT)

Peer-to-peer EV charger marketplace — hosts list chargers with availability windows; drivers discover and book nearby sessions via geo-coordinate proximity search. Built multi-strategy social + JWT authentication.

PrintIT — Cloud Print Broker (Nov 2022 – Mar 2023)

NestJS, MySQL, Google Drive API, IPP Protocol

On-the-go document printing via QR-scanned public printers. Used the Google Drive API as a zero-cost document-to-PDF conversion engine, IPP protocol for driverless printer communication, and ephemeral (auto-deleted) file storage for user document privacy.

Stargate — Subscription-Gated Crypto Community (Jun 2023)

NestJS, MySQL, TypeORM, MikroORM, Stripe, Google Calendar API

Cryptocurrency community platform with courses, trading signals, and portfolio tracking behind tiered Stripe subscriptions. Integrated Google Calendar API (via raw Axios) for bookable 1-on-1 analyst sessions.

EDUCATION & LANGUAGES

Bachelor of Science in Computer Science

[University of Lahore — Lahore, Pakistan](#)

August 2018 – June 2022

Languages: English (Proficient) | Urdu (Native)