Nick Mazurkin — Software Engineer, Team Lead, Staff Engineer

New York City, USA, greencard, +13477203170, career@mazurk.in, LinkedIn, GitHub, Resume

Java, **Python**, **Web**, **DB**, **No-SQL**, **Linux**, **ML**, **AI**, **Cloud**, **K8S**

PulsePoint, New York, USA

PulsePoint.com, AdTech, OpenRTB Exchange, DSP, DMP

Software Engineer with Expertise in Java, Python, and ML (Feb 2017 — present)

Experienced software engineer with a proven track record of developing and deploying high-performance applications using Java, Python, ML, Linux, and web technologies. Expertise in building and scaling distributed systems, implementing machine learning models, and optimizing performance for large-scale data processing.

- Developed an asynchronous web crawler/fetcher for /ads.txt and /app-ads.txt web resources, providing and enhancing transparency and trust for the SSP and DSP partners of our Exchange platform (processing ~500K domains daily, Java, Spring, Netty).
- Developed a supply prediction system for RTB traffic, enabling clients to switch to a cheaper self-serving platform (processing 24M documents daily using Lucene, Spring Framework and ML/CatBoost).
- Implemented a scoring model to effectively throttle proxy processing of incoming supply, selecting only 20% of the most profitable RTB requests and blocking 80% of the least profitable supply. This resulted in a 50% reduction in the number of required servers.
- Designed and implemented a DMP platform using asynchronous Java, gRPC, Cassandra, and Redis. This platform provides the effective solutions for targeting and optimization algorithms, accessing 30 billion entities with a p99 latency of 5ms and a p50 latency of 1ms.
- Developed a standalone service for domain attribution, enabling efficient domain and application bundle targeting (achieving 100k/sec QPS using gRPC and REST).
- Developed a client-side balancing library using the JSQ algorithm and circuit-breakers, significantly reducing client-server communication latency from p95=6ms to p95=1ms.
- Containerized and deployed a cluster of DoubleVerify service instances, achieving 300k/sec QPS in peak. Implemented custom client-side balancing and a low-latency custom client-to-cluster connection HTTP/REST library, enabling faster and easier deployments with efficient access.
- Developed a dockerized test infrastructure for custom configuration of Redis and Cassandra, similar to TestContainers, which facilitated the execution of integration tests.
- Developed a hierarchy of docker images for ~20 projects written in Python, Java, and Jupyter, significantly reducing the time required to create new projects.
- Developed an audience builder for 15 audiences, enabling our client to target very specific audiences with focus either on volume or accuracy (using Luigi, Hive, and ETL).
- Successfully resolved 2-3 support issues daily as part of a periodical 3rd-line technical on-call support rotation.
- Implemented several dozen RTB changes in the application stack, including improvements, new features, deployments, fixes, and monitoring (using Java, Netty, and Linux).
- Developed an engine for internal reports with embedded alerting, enabling transparent and quick monitoring of important pipelines and services (using Python, Hive, Impala, and Jupyter). Created 6 internal engineering reports.
- Implemented 7 ETL pipelines in Python using Luigi, Docker, Hive, and Presto.
- Improved video-completion rate model by making it more accurate and stable (Python, Scikit-learn)
- Implemented and delivered to production the look-alike audience model (Python, Catboost, Luigi, K8S)
- Developed a model to predict the quality of NPI attribution, enabling clients to select the confidence level for HCP targeting.
- Developed a model to predict the quality of ICD10 attribution based on obfuscated and aggregated clean-room prescription data, enabling clients to select DTC audiences based on condition probability.
- Developed a secure facade to access and process obfuscated clean-room data, providing a secure and efficient interface to private data (Python, Flask, Gunicorn, Google Cloud API).
- Implemented and delivered the gender prediction model (Python, Sklearn)

Team Lead, Data Science and Engineering team (Apr 2020 — present)

Led a team of 6 data scientists and 3 Java/Python engineers (DSE) working across 5 time zones, responsible for planning, performance tracking, coordination, and agile management. Proven ability to lead and mentor teams, manage projects, and deliver results on time and within budget.

- Successfully delivered 3 new internal services and 5 new ML models as a team leader.
- Successfully completed 20 quarterly plans with the team in research, development, and deployment.
- Successfully migrated DSE projects to the new data center.
- Provided interviewing, on-boarding, and mentoring for new team members.
- Successfully coordinated and communicated with other teams and team leaders on numerous projects.
- Successfully maintained an iterative Agile framework, including scrum, retrospectives, and prioritization of tasks.

Deutsche Bank

Deutsche Bank Technology Center, Moscow, RU

Senior Software Java Engineer, AVP (Jun 2016 — Feb 2017)

- Implemented several new algorithms, code and concurrency improvements to the dbCAP project achieving the required performance of the solution (exchange events messaging on Solace System message broker, KDB+, java).
- Successfully implemented the test suite for the legacy projects allowing to extend the life-time and reliability of the legacy

projects (process control, runtime log analyzers, TeamCity, environment settings).
- Improved the internal messaging framework based on Solace Systems message broker (achieved up to 160000 messages/sec stored reliably to the disk), introduced solutions for the performance metering.
- Designed and implemented the multithreaded and garbage-less algorithms and structures (Java) for message acknowledgement in complex environments to provide effective processing without long GC pauses.
- Implemented a modern front-end build toolchain for the legacy project, added the required UI features for the project (Angular 2, NPM, RxJs, TypeScript, TSLint, System.js, Rollup, jQuery, Bootstrap).

OK.ru

OK.ru, Moscow, RU, the one of the largest social networks (54 MAU)

Senior Software Java Developer (Oct 2012 — May 2016)

- Implemented a multiple features in the searching infrastructure of the OK.ru service: backend services, data storages, relevancy customization, indexation, replication (Apache Lucene, Spring Framework, Web Services, SQL/noSQL, Linux)
- Fixed multiple legacy issues in L12N infrastructure for the OK.ru service including the resources editor, Java frameworks and workflow process, achieved resource propagation to ~2000 servers.
- Make significant contribution to the moderation and antispam services, architecture and implementation, implemented NSFW detector, spam detector (based on naive bayes detector), UGC filter cluster (near real-time voting-quorum service).
- Successfully designed, implemented and run the MVP for the startup project (Tinder-alike) in very short time (3 months), created the architecture of the service including SOA, custom search service, asynchronous binary client protocol, push service (GCM), photo uploading/processing/distribution service, statistics and analytics (apache netty, apache cassandra, kafka, queues, messaging, mobile apps, python, machine learning).

Education

Mari State Technical University (1994 — 1999)

Yoshkar-Ola, Russia, “Computing machines, complexes, systems, and networks”, Engineer (with honors, equivalent to a master’s degree)

Open Source projects

net-crusher (2016)

TCP/UDP proxy library for automated and manual connectivity testing (used by RabbitMQ, Aerospike, Terracotta), more: project, sources, usages

Resume

The latest and the most detailed resume can be downloaded in different document formats at https://mazurk.in