ALEXEY SHISHKIN

+7(911)224-23-41 ⋄ alexxxeyshishkin@gmail.com ⋄ https://github.com/Alexvsalexvsalex

EDUCATION

HSE University   *2022 – 2024*
Master's degree
Applied Mathematics and Computer Science, bias in Computational Biology and Bioinformatics

ITMO University   *2018 – 2022*
Bachelor's degree
Applied Mathematics and Computer Science

SKILLS

Scala, Java, Python, C++, SQL, gRPC, sbt, bash, git, algorithms, highload, databases, MapReduce

WORK EXPERIENCE

Yandex. Spark over YTsaurus   June 2021 – Present
Software Engineer   *Scala, Java, Python, C++*

· Integration Apache Spark to YTsaurus – distributed storage and compute platform:

  – Designed architecture and implemented horizontally scalable and fault-tolerant Spark SQL server in C++ based on Apache Livy with patched bytecode, including managing users' sessions feature for quick reconnection;

  – Reduced query execution time by 30% by implementing filter predicate pushdown and extra plan optimization rules for joins, aggregations on sorted input data;

  – Supported streaming: checkpoint storage over YTsaurus, Kafka-like YTsaurus tables as source and sink;

  – Integrated TCP proxy routing NodePort traffic in Kubernetes installation;

  – Implemented allocation of ad-hoc RPC proxy over Unix socket inside user's container for big data processing;

  – Developed end-to-end and integration testing frameworks using ScalaTest and PyTest;

  – Prepared open-source release, configured Github Actions relying on S3 storage with publishing to Maven, PyPi and Docker registry.

ITMO University   Sept. 2021 – May 2022
Teaching assistant in the course Programming paradigms   *Java, JS, Clojure, Prolog*

· Review students' homework assignments and answer their questions about the course

PROJECTS

Dota 2 predictor   *Python*

· Telegram bot predicting results of live Dota 2 matches. The application includes a few parts:

  – Concurrent downloading match history from Valve API with InfluxDB metric storage;

  – Async auto-rebuilding CatBoost ML model with relevant dataset using Apache Kafka pipeline;

  – Monitoring streaming platform by Selenium and parsing live game state using computer vision.

ACHIEVEMENTS

Saint HighLoad++ 2023 Conference   Speaker
Google Hash Code 2022 Qualification Round   164th place (1.6%)
ICPC, quarter-final (2021)   23th place (26%)
All-Russian Olympiad of School Students, regional stage (2017, 2018)   1st place