

Dávid Szakállas

SRE & Platform

 Prague, Czechia

 +420 731 886 656

 david.szakallas@gmail.com

 dszakallas  szakallas.net

 dszakallas

About

Dávid is an experienced engineer specializing in developing and operating highly available systems, cloud platforms, and data pipelines on public clouds and Kubernetes. With a strong background in site reliability and distributed systems, he delivers robust infrastructure, observability, and automation. He holds an M.Sc. in Computer Science Engineering.

Skills

Software Engineering

Python	5.5 y
Scala	5.3 y
Go	3.8 y
Clojure	3.0 y

Site Reliability

CloudOps & Reliability	4.4 y
Observability (Prometheus, ELK)	3.3 y
Incident Response	2.4 y
CloudSec & NetSec (SPIFFE)	2.4 y

DevOps & Platform

Kubernetes (Apps & Platform)	7.3 y
Linux Containers	7.3 y
GitOps (Flux, Argo)	5.8 y
IaC & Terraform	4.4 y
Nix/NixOS	2.0 y

Public Clouds

AWS	5.3 y
Azure	4.9 y

Data Engineering

Apache Spark	6.4 y
Apache Airflow	4.4 y
AWS EMR & Data Lakes	4.0 y
Databricks	3.5 y

Misc

Citizenship
EU (Hungarian)

Languages
Hungarian (Native), English (Professional), Czech (Basic), German (Basic)

EXPERIENCE

Everpure (formerly Pure Storage)

Apr 2024 – Present

Staff Site Reliability Engineer

Developing and operating Everpure Cloud Azure Native. Defining SLOs/SLAs, region provisioning, building observability and alerting pipelines, incident response, and internal developer tooling.

azure k8s go gitops iac python terraform ELK stack SPIFFE vcluster github ir cloudsec cloudops devops netsec observability

Everpure (formerly Pure Storage)

Nov 2022 – Apr 2024

Staff Software Engineer

Developed backend services and workflows for SaaS/PaaS control planes written in Go and running on Kubernetes.

go k8s azure gitops backend engineering kubernetes operator development

Turbine

May 2021 – June 2022

Staff Software / Data Engineer

Evaluated and led adoption of Kubernetes (AKS) and GitOps (Flux). Productionized Apache Spark ETL workloads on Databricks with CI/CD and unit tests, and performance tuned PySpark applications.

k8s gitops python azure containers airflow databricks data lakes data engineering platform engineering

Graph Data Council

Oct 2020 – Sep 2022

Open Source Developer (Part-time contract)

Contributed to LDDB's Social Network Benchmark data generator in Scala and Apache Spark, optimizing generation of multi-terabyte datasets on EMR clusters.

aws emr apache spark apache hadoop scala database benchmarking

Hiya

June 2020 – Apr 2021

Senior Data Engineer

Developed and operated data lake on AWS Glue, Spark, Airflow, and Kubernetes. Automated Spark releases, built Prometheus metrics collection, and developed in-house libraries in Scala and Python.

scala apache spark aws gitops prometheus software engineering data lake design data platforms ops observability

Ekata

Oct 2017 – Feb 2020

Data Engineer

Developed the batch data ingestion pipeline for proprietary social graph database in Apache Spark, cutting ingestion time by 80%. Orchestrated supervised ML pipelines with Airflow, Kubernetes, and Databricks.

scala k8s airflow xgboost aws databricks software engineering graph databases ML training and inference data cleaning anomaly detection

EDUCATION

Budapest Univ. of Technology & Economics

2015 – 2018

M.Sc. in Computer Software Engineering

Thesis: openCypher graph query evaluation with a local search-based algorithm (in Clojure)

Budapest Univ. of Technology & Economics

2011 – 2015

B.Sc. in Computer Software Engineering

PROJECTS

Hobby & Open Source

-  mixxx-launchpad
-  bikeshed
-  podcaster
- Home Server
- Retrogaming

HID mappings for open-source DJ software (Mixxx)  
Nix utilities for workstation configuration
Podcast generation with Google NotebookLM  
Self-hosted media cluster on NixOS & k3s    
NixOS modules for emulation on Raspberry Pis   

Publications & Talks

2023
Oct 2018

LDDB: Driving competition and collaboration in graph data
Spark+AI Summit EU London Spark Schema for Free