Education

Sep 2022 - Feb 2025

ETH Zurich - MSc. Computer Science (GPA 5.71 / 6.0)

Master Thesis – "<u>Towards resource and interference-aware scheduling of ML workloads on GPUs</u>".
Wrote an ML cluster inference scheduler to improve GPU utilization by spatially colocating ML workloads on GPUs based on their resource requirements and interference patterns. (Efficient Architectures and Systems Lab, supervisor: Prof. Dr. Ana Klimovic)

Sep 2017 - Sep 2020

ETH Zurich - BSc. Computer Science (GPA: 5.4 / 6.0)

Bachelor Thesis – "Bomb the Internet, Infrastructure centralization and its weaknesses".
Simulated and analyzed the Internet's undergoing process of centralization and its robustness to large scale failures on the physical layer. (Network Security Group ETH Zurich, supervisor: Prof. Dr. Adrian Perrig)

Work Experience

Oct 2024 - Jan 2025

ETH Zurich - Research Assistant, Zurich

 Addressing the underutilization of GPUs by spatially co-locating multiple workloads in an interferenceaware manner with a particular focus on ML inference. (Sytems Group ETH Zurich, Prof. Dr Ana Klimovic)

Mar 2021 - Apr 2023

Archlet - Data Engineer, Zurich (full time alongside studies)

- Part of Archlet's backend team and responsible for the integration, extension and monitoring of several Python based microservices hosted as part of a Kubernetes cluster in Azure.
- Contributed to the development of Archlet's data pipelines and the operation of its optimization tool backed by mixed integer programming algorithms.
- Contributed to the technical transition into an event-driven microservice application using infrastructure as code.

Sep 2020 - Feb 2021

Oracle Labs - Research Assistant, Zurich

 Improved the security of Java applications running in containers, using static analysis to automatically generate SELinux policies at application and function granularity. (led to US patent application)

Publications and Patents

EuroMLSys '24

Foteini Strati, Paul Elvinger, Tolga Kerimoglu, and Ana Klimovic. 2024.

"ML Training with Cloud GPU Shortages: Is Cross-Region the Answer?". Proceedings of the 4th Workshop on Machine Learning and Systems. April 22, 2024, Athens, Greece. ACM, New York, NY, USA, 10 pages.

US20230074580A1

Neugschwandtner Matthias, Hugo Guiroux, and **Paul Elvinger**. "Runtime-tailored security policies for applications." U.S. Patent Application No. 17/465,694 filed March 9, 2023.

Selected Projects

Sep 2023 - Jan 2024

Efficient SDDMM Algorithms on GPU - github.com/francois141/dphpc

 Contributed to writing an efficient CUDA SDDMM implementation using adaptive tiling that outperforms Torch's implementation on large matrices.

Teaching

Spring 2024

Cloud Computing Architecture, ETH Zurich (Prof Dr. Ana Klimovic)

Fall 2018

Discrete Mathematics, ETH Zurich (Prof Dr. Ueli Maurer)

Programming Skills

Languages

Experienced: Python Intermediate: Java, C++, SQL (Postgres), Shell Scripting Basic: C, TypeScript

Dev tools + Frameworks

CUDA, Cloud Computing, Docker, Kubernetes, PyTorch, Linear Programming, Git, Unix, Vim

Cloud Infrastructure

Google Cloud Platform, Microsoft Azure

Languages

English, German, French – Full working proficiency; Luxembourgish – Mother tongue

Extra Curricula

Voluntary Work

Scout Group Leader - Daereldeieren Bridel (FNEL), Luxembourg

Organized weekly meetings for up to 50 children aged from 9 to 13 years.

Sport

Alpine Skiing

- Luxembourg National Champion in Slalom 2017
- Luxembourg Athlete at European Youth Olympic Festival (EYOF) 2015 in Austria

Golf (Hcp 6.6)

• Luxembourg Team Champion 2015, 2016, 2017

Marathon Runner