PEDRO F. SILVESTRE

(+351) 916272368 ◇ p.silvestre21@imperial.ac.uk ◇ 30/06/1997 ◇ Portuguese
pfsilvestre.me ◇ PSilvestre ◇ in/pedro-silvestre ◇ pmfsilvestre

EDUCATION

PhD in Dataflow Systems *Feb 2021 - Present*
Imperial College London, London
Large Scale Data & Systems Group — Advisors: P. Pietzuch & H. Pirk

- Research at the intersection of Stream Processing, Database and Deep Learning systems.
- Currently designing a solution to scale-out and improve the usability of Deep Learning compilers.

MSc in Computer Science *Sep 2018 - Dec 2020*
NOVA School of Science and Engineering, Lisbon — Grade Average: 18/20
Department of Informatics — Advisors: A. Katsifodimos & J. Leitão
Thesis: *Clonos: Consistent High-Availability for Distributed Stream Processing through Causal Logging*

Notable Project: (Concurrency and Parallelism) Developed and evaluated a C library providing work-efficient implementations of common parallel programming patterns using Cilk. Devised a variation of the Blelloch scan which accepts any input size. *Project Grade: 19/20*

Exchange Semester *Feb 2019 - Jul 2019*
Delft University of Technology, Delft — Grade Average: 9/10
Faculty of Electrical Engineering, Mathematics & Computer Science

Notable Project: (Web-scale Data Management) Implemented an indefinitely scalable service using Akka Cluster in a microservice architecture. Implemented SAGAs for distributed transactions and Event Sourcing used for fault-tolerance. Dynamic scaling using Auto-Scaling Groups. *Project Grade: 9.5/10*

BSc in Computer Science *Sep 2015 - Jul 2018*
NOVA School of Science and Engineering, Lisbon — Grade Average: 17/20
Department of Informatics

Notable Project: (Distributed Systems) Built an HDFS clone with Namenodes and Datanodes. Ring replication was used for data fault-tolerance. Service discovery via Kafka or multicast communication. A functioning Map-Reduce engine was also implemented. *Project Grade: 20/20*

RESEARCH EXPERIENCE

Research Engineer *Jun 2019 - Nov 2020*
Delft University of Technology, Delft
Web Information Systems Group

- Led the design and implementation of Clonos (`delftdata.github.io/clonos-web`), a Stream Processing System using Causal Logging for consistent local recovery and high-availability.
- Built custom infrastructure for virtual machine deployment in the Dutch *SurfSara* cluster.
- Developed automated distributed benchmarking infrastructure for Stream Processors by leveraging Kubernetes, capturing real-time end-to-end throughput, latency and recovery time with millisecond precision.
- Participated in the design, development and testing of rho (ρ), a stateful FaaS platform. Built tooling for the authoring and deployment of stateful functions.

Research Assistant *Sep 2018 - Dec 2018*
NOVA School of Science and Engineering, Lisbon
NOVA-LINCS Research Laboratory

- Implemented a middleware layer providing transparent δ-CRDT based state synchronization for wireless AdHoc sensor networks in C. A reliable message fragmentation protocol was also added.

PUBLICATIONS

SIGMOD'21
(ranked Core A*)

Clonos: Consistent Causal Recovery for Highly-Available Streaming Dataflows.
Silvestre, P. F., Fragkoulis, M., Spinellis, D., & Katsifodimos, A. (2021, June).
In Proceedings of the 2021 International Conference on Management of Data (pp. 1637-1650).

PROFESSIONAL EXPERIENCE

Big Data Software Engineering Internship *Jul 2018 - Sep 2018*
XPandIT, Lisbon

- Full-stack development of a web application for orchestrating Docker containers for data-science workloads, integrating with Kerberos for single sign-on into containers. Containers were automatically built from a web form describing the tools and resources the container should have.

(Academic) Software Engineering & Quality Assurance Internship *Mar 2018 - July 2018*
Feedzai, Lisbon Grade: 19/20

- Deployed Kubernetes in the on-premises cluster. Deployed a CI solution (Jenkins) with dynamic executor provisioning on Kubernetes cluster, improving CI resource usage by up to 30%.
- Achieved elasticity by joining AWS EC2 instances dynamically to Kubernetes automatically.
- Modified internal integration testing libraries to request resources from Kubernetes cluster.
- Presented the solution to over 100 colleagues during internal talks.

HONORS & ACHIEVEMENTS

Winner of the HackDelft 2019 Hackathon (40 teams)
Project: Built an early warning anomaly detection system for the Dutch railroad network which processed time series sensor data in real time. Warnings were presented in a web application which included automated visualization of abnormal sensor data.

Awarded 1^{st} prize in CLC Merit Scholarship (€5000)
Awarded the CM Azambuja Merit Scholarship (€1000) x4

PROJECTS

Process Controller Simulator: A highly flexible simulation framework implemented in Python. Able to concurrently simulate complex chemical processes, controllers (e.g. MPC) and more. Includes a web interface for creating and visualizing simulations. I am using the project to teach a chemical engineering student Python.

Raspberry Pi Cluster: Assembled a 4 node cluster with compact power and ethernet delivery. Runs Kubernetes and Slurm (for OpenMPI) on top of which I deploy services such as Jenkins and personal websites.

OTHER HIGHLIGHTS

Volunteering and Presenting at Conferences: Volunteered at *Flink Forward 2019*, where I also presented the *rho* (ρ) project. Presented *Clonos*, virtually, at *SIGMOD'21*, where I learned useful video editing skills.

Object Oriented Programming Tutor: For a semester, tutored 2 undergraduate students in OOP. Prepared lectures, which were later presented and provided guidance in projects.

Student Worker: For a year, managed my time between work and studies, in different countries, whilst maintaining good performance in both, demonstrating solid organizational and time management skills.

LANGUAGES

Portuguese	Native Proficiency
English	Full Professional Proficiency (IELTS: 8.5/9, CEFR level C2)
Spanish	Limited Working Proficiency

References - available upon request