Ricardo Macedo

Assistant Researcher at INESC TEC · Invited Assistant Professor at University of Minho

Dept. Informática, Universidade do Minho, Braga, Portugal (+351) 253 604 477 ricardo.g.macedo@inesctec.pt
rgmacedo.github.io rgmacedo Google Scholar ORCID DBLP

Summary

Assistant Researcher at INESC TEC (Distributed Storage Research team) and Invited Assistant Professor at the University of Minho. My research lies in the areas of storage and operating systems, focusing on designing new building blocks fitted for the performance, reliability, and sustainability requirements of modern large-scale infrastructures, namely cloud computing data centers and HPC supercomputers.

Personal Information

Name	Ricardo Gonçalves Macedo
Inst. Page	https://www.inesctec.pt/en/people/ricardo-goncalves-macedo
Pers. Page	https://rgmacedo.github.io/
ORCID	0000-0003-4036-0126
Google Scholar	3QPFbOcAAAAJ
DBLP	159/9543
ACM	99659535288
GitHub	rgmacedo

Current position

2023 – cur. **Assistant Researcher**, *INESC TEC*.

Research in storage and operating systems fitted for the performance, reliability, and sustainability of modern large-scale infrastructures.

Leading the *"Storage and Operating Systems for Modern Data Centers"* research line at INESC TEC's *Distributed Storage Research* team.

Scientific coordinator of INESC TEC's computational infrastructures *CLOUDinha Lab*.

2020 – cur. **Invited Assistant Professor (25%)**, *University of Minho*.

Teaching *Distributed Systems* (Bachelor's in Informatics Engineering, LEI, 3rd year; 1st semester), and *Data Storage Systems* and *Advanced Computing Projects* (Master's in Informatics Engineering, MEI, 1st year; 2nd semester).

Education

2017 – 2023 **MAP-i Doctoral Program in Computer Science.**
Universities of Minho, Aveiro, and Porto, Portugal.
Advisors: João Tiago Paulo and José Orlando Pereira.
Thesis: *User-level Software-Defined Storage Data Planes*.
Classification: Very good (UMinho's highest grade).

2011 – 2017 **Integrated Master's in Informatics Engineering.**
University of Minho, Portugal.
Advisors: João Tiago Paulo and Rui Carlos Oliveira.
Thesis: *Secure Computation in NoSQL Databases*.
Classification (thesis): 20 out of 20.

Experience

2017 – 2023 **Researcher (PhD student)**, *HASLab, INESC TEC*.
Research in Software-Defined Storage, local and distributed storage, key-value storage, kernel-bypass storage.

2020 (summer) **Research Intern**, *AIST*, Tsukuba, Japan (remote).
Advisors: Jason Haga and Yusuke Tanimura.
Research in storage data planes for ensuring bandwidth guarantees in the ABCI supercomputer.

2016 – 2017 **Researcher (MSc student)**, *HASLab, INESC TEC*.
Research in secure computation over NoSQL and SQL databases.

Publications

Conferences

EDBT 2027 Olivier Michaud, **Ricardo Macedo**, Oana Balmau, Bettina Kemme. Query-Aware DVFS For GPU-Accelerated Relational Databases. In *30th International Conference on Extending Database Technology*, 2027 (to appear).

EuroSys 2026 Rahma Nouaji, Stella Bitchebe, **Ricardo Macedo**, Oana Balmau. MinatoLoader: Accelerating Machine Learning Training Through Efficient Data Preprocessing. In *21st ACM European Conference on Computer Systems*, 2026. **Best Poster Award.**

ICPE 2026 José Pedro Peixoto, Alexis Gonzalez, Janki Bhimani, Raju Rangaswami, Cláudia Brito, João Paulo, **Ricardo Macedo**. Holpaca: Holistic and Adaptable Cache Management for Shared Environments. In *17th ACM/SPEC International Conference on Performance Engineering*, 2026.

Middleware 2026 Mariana Miranda, Yusuke Tanimura, Jason Haga, Amit Ruhela, Stephen Lien Harrell, John Cazes, José Pereira, **Ricardo Macedo**, João Paulo. A Tale of Scale: Enabling Scalable and Holistic Storage QoS for Exascale HPC Systems. In *27th ACM International Middleware Conference*, 2026.

SBAC-PAD 2026 Fábio Souza, Daniel Sodré, João Paulo, **Ricardo Macedo**, Cristina Boeres, Vinod Rebello, Felipe Portella, Paulo Estrela, Renzo Malini. Phoning Home: Repatriating output from cloud jobs with a user-level, provider-agnostic data transfer tool. In *38th IEEE/SBC International Symposium on Computer Architecture and High Performance Computing*, 2026.

VLDB 2025[2] Rúben Adão, Zhongjie Wu, Changjun Zhou, Oana Balmau, João Paulo, **Ricardo Macedo**. Keigo: Co-designing Log-Structured Merge Key-Value Stores with a Non-Volatile, Concurrency-aware Storage Hierarchy. In *51st International Conference on Very Large Data Bases*, 2025.

VLDB 2024[1] Maria Ramos, João Azevedo, Kyle Kingsbury, José Pereira, Tânia Esteves, **Ricardo Macedo**, João Paulo. When Amnesia Strikes: Understanding and Reproducing Data Loss Bugs with Fault Injection. In *50th International Conference on Very Large Data Bases*, 2024.

CCGrid 2023 **Ricardo Macedo**, Mariana Miranda, Yusuke Tanimura, Jason Haga, Amit Ruhela, Stephen Lien Harrell, Richard Todd Evans, José Pereira, João Paulo. Taming Metadata-intensive HPC Jobs Through Dynamic, Application-agnostic QoS Control. In *23rd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing*, 2023.

CCGrid 2022 Marco Dantas, Diogo Leitão, Peter Cui, **Ricardo Macedo**, Xinlian Liu, Weijia Xu, João Paulo. Accelerating Deep Learning Training Through Transparent Storage Tiering. In *22nd IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing*, 2022.

FAST 2022 **Ricardo Macedo**, Yusuke Tanimura, Jason Haga, Vijay Chidambaram, José Pereira, João Paulo. PAIO: General, Portable I/O Optimizations With Minor Application Modifications. In *20th USENIX Conference on File and Storage Technologies*, 2022.

SYSTOR 2021 Alberto Faria, **Ricardo Macedo**, José Pereira, João Paulo. BDUS: Implementing Block Device Drivers in User Space. In *14th ACM International System and Storage Conference*, 2021. **Best Paper Runner-up.**

SRDS 2017 **Ricardo Macedo**, João Paulo, Rogério Pontes, Bernardo Portela, Tiago Oliveira, Miguel Matos, Rui Oliveira. A Practical Framework for Privacy-Preserving NoSQL Databases. In *36th IEEE International Symposium on Reliable Distributed Systems*, 2017.

[1], [2]: the same paper, published in PVLDB (listed under Journals) and presented at the VLDB conference.

Journals

TACO 2026 Manoj Saha, Yanzhao Wu, Cláudia Brito, Raju Rangaswami, João Paulo, **Ricardo Macedo**, Janki Bhimani. Fragment: Efficient DNN Checkpoint with Related Model Consistency. In *ACM Transactions on Architecture and Code Optimization, 23 (3)*, 2026.

PVLDB 2025[2] Rúben Adão, Zhongjie Wu, Changjun Zhou, Oana Balmau, João Paulo, **Ricardo Macedo**. Keigo: Co-designing Log-Structured Merge Key-Value Stores with a Non-Volatile, Concurrency-aware Storage Hierarchy. In *Proceedings of the VLDB Endowment, 18 (9)*, 2025.

PVLDB 2024[1] Maria Ramos, João Azevedo, Kyle Kingsbury, José Pereira, Tânia Esteves, **Ricardo Macedo**, João Paulo. When Amnesia Strikes: Understanding and Reproducing Data Loss Bugs with Fault Injection. In *Proceedings of the VLDB Endowment, 17 (11)*, 2024.

Access 2023 Tânia Esteves, **Ricardo Macedo**, Rui Oliveira, João Paulo. Toward a practical and timely diagnosis of applications' I/O behavior. In *IEEE Access (11)*, 2023.

CSUR 2020 **Ricardo Macedo**, João Paulo, José Pereira, Alysson Bessani. A Survey and Classification of Software-Defined Storage Systems. In *ACM Computing Surveys 53, 3 (48)*, 2020.

Workshops

SC-W 2026 Alexandre Fernandes, Alícia Oliveira, Beatriz Cepa, Cláudia Brito, **Ricardo Macedo**. Watts behind the weights: Characterizing energy consumption of the AI training lifecycle in HPC. In *Supercomputing Workshops: AI on HPC*, co-located with International Conference for High Performance Computing, Networking, Storage, and Analysis, 2026.

RLDM 2025 José Peixoto, Alexis Gonzalez, Janki Bhimani, Raju Rangaswami, Cláudia Brito, João Paulo, **Ricardo Macedo**. Idiosyncrasies of Programmable Caching Engines. In *Workshop on Reliable Large-scale Data Management*, co-located with IEEE International Symposium on Reliable Distributed Systems, 2025.

RLDM 2025 André Ferreira, Gonçalo Sousa, Amit Ruhela, John Cazes, Janki Bhimani, Raju Rangaswami, **Ricardo Macedo**, Cláudia Brito, João Paulo. One Does Not Simply Cache: A Study on Distributed Caches for AI Training in HPC. In *Workshop on Reliable Large-scale Data Management*, co-located with IEEE International Symposium on Reliable Distributed Systems, 2025.

SC-W 2024 Mariana Miranda, Yusuke Tanimura, Jason Haga, Amit Ruhela, Stephen Lien Harrell, John Cazes, **Ricardo Macedo**, José Pereira, João Paulo. Can Current SDS Controllers Scale to Modern HPC Infrastructures? In *Supercomputing Workshops: Communication, I/O, and Storage at Scale on Next-Generation Platforms*, co-located with International Conference for High Performance Computing, Networking, Storage, and Analysis, 2024.

DCDS 2023 Tânia Esteves, **Ricardo Macedo**, Rui Oliveira, João Paulo. Diagnosing applications' I/O behavior through system call observability. In *5th Workshop on Data-Centric Dependability and Security*, co-located with IEEE/IFIP International Conference on Dependable Systems and Networks, 2023.

REX-IO 2022 **Ricardo Macedo**, Mariana Miranda, Yusuke Tanimura, Jason Haga, Amit Ruhela, Stephen Lien Harrell, Richard Todd Evans, João Paulo. Protecting Metadata Servers From Harm Through Application-level I/O Control. In *2nd Workshop on Re-envisioning Extreme-Scale I/O for Emerging Hybrid HPC Workloads*, co-located with IEEE International Conference in Cluster Computing, 2022.

WoC 2021 Alberto Faria, **Ricardo Macedo**, João Paulo. Pods-as-Volumes: Effortlessly Integrating Storage Systems and Middleware into Kubernetes. In *7th International Workshop on Container Technologies and Container Clouds*, co-located with ACM/IFIP International Middleware Conference, 2021.

REX-IO 2021 **Ricardo Macedo**, Cláudia Correia, Marco Dantas, Cláudia Brito, Weijia Xu, Yusuke Tanimura, Jason Haga, João Paulo. The Case for Storage Optimization Decoupling in Deep Learning Frameworks. In *1st Workshop on Re-envisioning Extreme-Scale I/O for Emerging Hybrid HPC Workloads*, co-located with IEEE International Conference in Cluster Computing, 2021.

REX-IO 2021 Marco Dantas, Diogo Leitão, Cláudia Correia, **Ricardo Macedo**, Weijia Xu, João Paulo. Monarch: Hierarchical Storage Management for Deep Learning Frameworks. In *1st Workshop on Re-envisioning Extreme-Scale I/O for Emerging Hybrid HPC Workloads*, co-located with IEEE International Conference in Cluster Computing, 2021.

SRDSW 2019 **Ricardo Macedo**, Alberto Faria, João Paulo, José Pereira. A Case for Dynamically Programmable Storage Background Tasks. In *38th International Symposium on Reliable Distributed Systems Workshops*, 2019.

SRDSW 2019 Tânia Esteves, **Ricardo Macedo**, Alberto Faria, Bernardo Portela, João Paulo, José Pereira, Danny Harnik. TrustFS: An SGX-enabled Stackable File System Framework. In *38th International Symposium on Reliable Distributed Systems Workshops*, 2019.

Works under submission

2026 Miguel Braga, Júlio Pinto, Rahma Nouaji, Olivier Michaud, Bettina Kemme, Oana Balmau, Cláudia Brito, **Ricardo Macedo**. A principled approach for energy-efficient training via phase-aware GPU frequency tuning. In *arXiv preprint arXiv:2609.24205*, 2026.

Preprints

2026 José Peixoto, Alexis Gonzalez, Janki Bhimani, Raju Rangaswami, Cláudia Brito, João Paulo, **Ricardo Macedo**. Idiosyncrasies of Programmable Caching Engines. In *arXiv preprint arXiv:2603.14357*, 2026. *Published at RLDM 2025.*

2025 Rahma Nouaji, Stella Bitchebe, **Ricardo Macedo**, Oana Balmau. MinatoLoader: Accelerating Machine Learning Training Through Efficient Data Preprocessing. In *arXiv preprint arXiv:2509.10712*, 2025. *Published at EuroSys 2026.*

Rúben Adão, Zhongjie Wu, Changjun Zhou, Oana Balmau, João Paulo, **Ricardo Macedo**. Keigo: Co-designing Log-Structured Merge Key-Value Stores with a Non-Volatile, Concurrency-aware Storage Hierarchy (Extended Version). In *arXiv preprint arXiv:2506.14630*, 2025. *Extended version of the VLDB 2025 / PVLDB paper.*

2023 Tânia Esteves, **Ricardo Macedo**, Rui Oliveira, João Paulo. Diagnosing applications' I/O behavior through system call observability. In *arXiv preprint arXiv:2304.08569*, 2023. *Published at DCDS 2023.*

Ricardo Macedo, Mariana Miranda, Yusuke Tanimura, Jason Haga, Amit Ruhela, Stephen Lien Harrell, Richard Todd Evans, José Pereira, João Paulo. PADLL: Taming Metadata-intensive HPC Jobs Through Dynamic, Application-agnostic QoS Control. In *arXiv preprint arXiv:2302.06418*, 2023. *Published at CCGrid 2023.*

2021 **Ricardo Macedo**, Yusuke Tanimura, Jason Haga, Vijay Chidambaram, José Pereira, João Paulo. PAIO: A Software-Defined Storage Data Plane Framework. In *arXiv preprint arXiv:2106.03617*, 2021. *Published at FAST 2022.*

Theses

PhD Thesis **Ricardo Macedo**. User-level Software-Defined Storage Data Planes. *PhD Thesis, Universities of Minho, Aveiro, and Porto*. Supervised by João Paulo and José Pereira. 2023.

MSc Thesis **Ricardo Macedo**. Computação Segura em Bases de Dados NoSQL. *Master's Thesis, Universidade do Minho*. Supervised by João Paulo and Rui Oliveira. 2017.

Projects

Ongoing

2026 – cur. **GenSEMICON: GENerative AI Chips and Software Engineering Stack: Integrating EU-Sovereign SEMICONductors and Edge Intelligence Systems**.

Researcher. Research on designing power-aware techniques for improving AI training energy-efficiency. *Partners:* INESC TEC, MAG, DNV, ICCS, ETS, TUC, NTNU, UPB, and others.

2025 – cur. **PATA: Power- and Thermal-aware Management for Energy-Efficient AI Training in HPC Infrastructures**.

Co-PI with John Cazes (TACC, UT Austin). Research on designing power- and thermal-aware techniques for improving AI training energy-efficiency. *Partners:* INESC TEC, TACC.

2025 – cur. **CDMS: Claim Denial Management Solution**.

Local PI. Research on designing new performance and energy-efficient primitives for LLM training on HPC infrastructures. *Partners:* INESC TEC, GetEffy, RandyLabs.

2024 – cur. **BCD.S+M: Sistema modular de Armazenamento e Gestão de Dados em Blockchain com IA**.

Researcher. Research on designing new data caching and tiering strategies for multi-storage, multi-device environments. *Partners:* INESC TEC, InvisibleLab.

2024 – cur. **CENTRA-PASM: Performance and Sustainability Management in Modern Data Centers**.

PI. Research on designing new building blocks fitted for the performance and sustainability requirements of modern, large-scale infrastructures. *Partners:* INESC TEC, AIST, FIU.

2018 – cur. **CENTRA-ESDM: Efficient and Secure Data Management for HPC and Cloud Computing**.

Researcher. Research and development of Software-Defined Storage systems to improve the programmability and configurability of HPC and Cloud Computing infrastructures.

Concluded

2025 – 2026 **DisaggregatedHPC: Towards energy-efficient, software-managed resource disaggregation in HPC infrastructures**.

Co-PI with Weicheng Huang (NCHC, NIAR). Research on designing a software-managed resource disaggregation framework for HPC infrastructures. *Partners:* INESC TEC, NIAR.

2021 – 2026 **Sustainable HPC**.

Researcher. Research on designing new energy efficient systems for measuring and controlling the energy consumption of HPC infrastructures. *Partners:* INESC TEC, INEGI.

2023 – 2026 **ATE: Alliance for Energy Transition**.

Researcher. Research on designing new systems for measuring and controlling the energy consumption of modern workloads under large-scale infrastructures. *Partners:* INESC TEC, EFACEC (more than 80 partners).

2020 – 2023 **BigHPC: A Management Framework for Consolidated Big Data and HPC**.

POCI-01-0247-FEDER-045924

Researcher. Research and development of Software-Defined Storage systems to improve performance isolation, QoS, and fairness in converged HPC and Big Data infrastructures.

2022 **LazyFS: Lazy File System Project**.

PP2022-0019

Researcher. Consulting services with Jepsen LLC. Research and development of fault-injection file system.

2020 – 2021 **PAStor: Programmable and Adaptable Storage for AI-oriented HPC Ecosystems**.

UTA-EXPL/CA/0075/2019

Researcher. Research and development of a Software-Defined Storage data plane for improving I/O performance of AI-powered applications in HPC infrastructures.

2018 – 2019 **IBM Research Haifa: Joint Study Agreement**.

Researcher. Research and development of a programmable and modular stackable file system framework for implementing secure content-aware storage functionalities over hardware-assisted trusted execution environments.

2016 – 2018 **SafeCloud**.

H2020-DS-2014-1/653884

Researcher. Design, development, and evaluation of a privacy-preserving framework for NoSQL databases. Integration with a SQL query engine in order to provide secure computation over real-time SQL-based applications.

Advising

PhD Students

Starts in 2027 **Alexandre Fernandes.** *Optimizing the Energy Efficiency of Modern Data Centers.* Co-advised with Cláudia Brito. Doctoral Programme in Informatics (PDInf), University of Minho.

Daniel Pereira. *Rethinking Data Management Systems for the Modern Era.* Doctoral Programme in Informatics (PDInf), University of Minho.

2026 – cur. **André Ferreira.** *Improving Quality-of-Service and Performance for HPC Storage of Exascale Supercomputers.* Co-advised with João Paulo. Doctoral Programme in Informatics (PDInf), University of Minho.

2025 – cur. **Fábio Souza.** *Hybrid Data Processing and Storage across Cloud and HPC Environments.* Co-advised with João Paulo. Doctoral Programme in Informatics (PDInf), University of Minho.

2024 – cur. **Pedro Peixoto.** *Software-Managed Resource Disaggregation for Large-Scale HPC Infrastructures.* Co-advised with João Paulo. Doctoral Programme in Informatics (PDInf), University of Minho.

Rahma Nouaji. *Systems Optimizations for Machine Learning.* Co-advised with Oana Balmau. Ph.D. in Computer Science, McGill University.

Rúben Adão. *Heterogeneous Storage Management in Modern I/O Infrastructures.* Co-advised with Oana Balmau and João Paulo. Doctoral Programme in Informatics (PDInf), University of Minho.

MSc Students

2026 – cur. **André Miranda.** *Energy monitoring systems for distributed applications.* Co-advised with Cláudia Brito and Alexandre Fernandes. University of Minho.

José Soares. *Storage optimizations for vector databases.* University of Minho.

2025 – cur. **Daniel Pereira.** *LegoKV: A key-value store for disaggregated environments.* Co-advised with Rúben Adão. University of Minho.

Duarte Ribeiro. *Strict durability in LSM-based key-value stores.* University of Minho.

Júlio Pinto. *Energy control system for large-scale AI training.* University of Minho.

Rui Lopes. *Generalized storage tiering strategies.* Co-advised with Miguel Matos. University of Minho.

2024 – 2025 **Alexandre Fernandes.** *Energy Monitoring System for Large-Scale Infrastructures.* Co-advised with Cláudia Brito. 19 out of 20. University of Minho.

Carlos Machado. *An NVMe-oriented I/O Scheduler for Reducing Energy Consumption.* Co-advised with Cláudia Brito. Ongoing. University of Minho.

Miguel Braga. *GPU Energy Control for LLM Models.* Co-advised with Cláudia Brito. 19 out of 20. University of Minho.

2023 – 2024 **Diana Rodrigues.** *GPU Energy Control for Deep Learning Systems.* Co-advised with Cláudia Brito and António Sousa. 19 out of 20. University of Minho.

José Pedro Fernandes. *Comprehensive Study of the Energy Impact of Key-value Stores.* Co-advised with Cláudia Brito and António Sousa. 19 out of 20. University of Minho.

Mariana Amorim. *Energy Control System for Disaggregated Storage Resources.* Co-advised with Cláudia Brito and João Paulo. 18 out of 20. University of Minho.

Sara Pereira. *Towards Fine-grained, Holistic Energy Control in Large-Scale Computing Infrastructures.* Co-advised with Carlos Baquero and Cláudia Brito. 18 out of 20. University of Porto.

2022 – 2023 **Maria Beatriz Moreira.** *Distributed I/O Optimizations for Deep Learning Training.* Co-advised with João Paulo and Cláudia Brito. 18 out of 20. University of Minho.

Maria Ramos. *Reproducible Fault Injection in Local Storage Systems.* Co-advised with João Paulo and Tânia Esteves. 19 out of 20. University of Minho.

Pedro Peixoto. *Towards Programmable and Adaptable Caches.* Co-advised with João Paulo. 19 out of 20. University of Minho.

Rúben Adão. *Co-designing Log-Structured Merge Key-Value Stores with a Non-Volatile Storage Hierarchy.* Co-advised with João Paulo. 19 out of 20. University of Minho.

2021 – 2022 **Alexandre Miranda.** *Realistic Assessment of Failures in the SPDK Platform.* Co-advised with João Paulo. 18 out of 20. University of Minho.

João Azevedo. *LazyFS: A file system for assessing applications data durability.* Co-advised with João Paulo. 20 out of 20. University of Minho.

2020 – 2021 **Marco Dantas.** *Accelerating Deep Learning Training on High-Performance Computing with Storage Tiering.* Co-advised with João Paulo and Rui Oliveira. 20 out of 20. University of Minho.

2019 – 2020 **Cláudia Mendonça.** *PRISMA: A Prefetching Storage Middleware for Accelerating Deep Learning Frameworks.* Co-advised with João Paulo and António Sousa. 18 out of 20. University of Minho.

Diogo Leitão. *RSafeFS: A Modular File System for Remote Storage.* Co-advised with João Paulo and José Orlando Pereira. 17 out of 20. University of Minho.

Researchers

2026 – cur. **Alexandre Fernandes.** *Energy-efficient systems for large-scale infrastructures.* Work within the CDMS and PATA projects. INESC TEC.

Undergraduate Research Projects

2024 **Diana Rodrigues**, **José Pedro Fernandes**, and **Mariana Amorim**. *Monitoring Energy Consumption with Fine-Granularity in Data-centric Applications.* Co-advised with António Sousa and Cláudia Brito. University of Minho.

Francisco Neves and **Guilherme Fernandes**. *Metadata Trace Replayer for Large-Scale HPC File Systems.* Co-advised with João Paulo. University of Minho.

2023 **Maria Beatriz Moreira**, **Pedro Peixoto**, and **Rúben Adão**. *Programmable and Adaptable Data Transformations.* Co-advised with João Paulo. University of Minho.

2022 **Susana Marques** and **Tomás Costa.** *A Key-Value Storage Benchmarking Tool.* Co-advised with João Paulo. University of Minho.

2021 **Alberto Faria.** *An Evaluation of Linux Interfaces for Storage I/O.* Co-advised with João Paulo. University of Minho.

Research Mentor

2024 – 2026 **Edgar Araújo**. *Comprehensive study of resource usage patterns in large-scale HPC infrastructures.* Co-advised with Pedro Peixoto.

2025 **Alexis Gonzalez**. *Rethinking storage systems for emerging I/O devices.* Co-advised with Raju Rangaswami and Pedro Peixoto.

Humberto Gomes. *Heterogeneous GPU scheduling for AI training workloads.* Co-advised with André Pereira.

2024 – 2025 **Daniel Pereira**. *LSM-based key-value stores for disaggregated storage.*

2024 **Rahma Nouaji**. *I/O optimizations for Large Language Models.* Co-advised with Oana Balmau and Cláudia Brito.

2022 – 2024 **Guilherme Fernandes**. *Metadata trace replayer for HPC parallel file systems.*

Diana Rodrigues, **João Fernandes**, **Mariana Amorim** and **Sara Pereira**. *Exploring energy-aware policies for modern I/O infrastructures.*

2021 – 2022 **Maria Beatriz Moreira** and **Pedro Peixoto.** *Building transparent data transformations mechanisms (compression, encryption).*

Rúben Adão. *Storage tiering strategies for LSM-based key-value stores.*

2021 **Diogo Ribeiro.** *Storage tiering strategies in FUSE-based file systems.*

Research Hosting

May – Jul 2025 **Alexis Gonzalez.** PhD student, FIU. *Rethinking storage systems for emerging I/O devices.* INESC TEC's Visiting Researcher Programme 2024.

May – Jul 2024 **Rahma Nouaji.** PhD student, McGill University. *I/O optimizations for large language models.* INESC TEC's Visiting Researcher Programme 2023.

Jun 2022 **Oana Balmau.** Assistant Professor, McGill University. *Storage tiering for LSM-based key-value stores.*

Teaching

2026 – cur. **Advanced Computing Projects**.

Invited Assistant Professor (25%).

Informatics Engineering Master's (1st year) and Masters in Advanced Computing (1st year), University of Minho.

2025 – cur. **Data Storage Systems**.

Invited Assistant Professor (25%).

Informatics Engineering Master's (1st year) and Masters in Advanced Computing (1st year), University of Minho.

2020 – cur. **Distributed Systems**.

Invited Assistant Professor (25%).

Bachelor's in Informatics Engineering (LEI, 3rd year; formerly Integrated Master's, MIEI), University of Minho.

2019 – 2025 **Operating Systems**.

Invited Assistant Professor (25%).

Bachelor's in Informatics Engineering (LEI, 2nd year; formerly Integrated Master's, MIEI), University of Minho.

Service

Organization

2027 CHEOPS Workshop (EuroSys) – PC Chair.

2026 PDSW Workshop (SC) – Reproducibility Chair.

2025 PDSW Workshop (SC) – Reproducibility Chair.

Program Committees

2027 EuroSys (Extra PC), IPDPS, ICS, ICPE, CCGrid, SCA, EuroSys (posters).

2026 ICS, ICPE, ICPP, CCGrid, CLUSTER, SBAC-PAD, HPDC (posters), and workshops at IPDPS (ESSA), EuroSys (InSyDe, CHEOPS), and HPDC (REX-IO).

2025 SSDBM, Euro-Par (posters and demos), and workshops at SRDS (RLDM), CLUSTER (REX-IO), and IPDPS (ESSA).

2024 ASPLOS, Euro-Par (posters and demos).

2023 EuroSys Shadow PC, and workshops at IPDPS (ESSA) and CLUSTER (REX-IO).

2022 REX-IO Workshop (CLUSTER).

Journal Reviewer

2026 Future Generation Computer Systems.

2025 ACM Transactions on Architecture and Code Optimization (TACO).

2024 ACM TACO, Journal of Supercomputing.

2023 Journal of Supercomputing.

External Reviewer

2022 DSN.

2021 DSN, ICDCS.

2019 OPODIS, SRDS, DSN.

2018 Middleware.

Other Service

2018 **Volunteer**, ACM European Conference on Computer Systems *(Porto, Portugal)*.

Awards, honors, and scholarships

As a researcher

2026 **Best Poster Award.** 21st ACM European Conference on Computer Systems (EuroSys 2026), for the paper *"MinatoLoader: Accelerating Machine Learning Training Through Efficient Data Preprocessing"*.

2026 **Best Junior Reviewer Award.** 17th ACM/SPEC International Conference on Performance Engineering.

2025 **EU Innovation Radar.** European Commission. Energy efficiency monitoring tool listed as a key innovation in the EU Innovation Radar program, within the *"Testing framework for AI energy efficiency"*.

2023 **Cor Baayen Early Career Researcher Award.** *Finalist.* European Consortium for Informatics and Mathematics (ERCIM).

As a student

2022 **"The Incredibles" Award.** *INESC TEC, Portugal.* Award given by INESC TEC and HASLab coordinators to João Paulo and Ricardo Macedo, due to their contribution to the field of storage, together with their respective teams of MSc students.

2021 **Best Paper Runner-up.** *"BDUS: Implementing Block Device Drivers in User Space"*, *14th ACM International System and Storage Conference.*

2021 **ATC'21 Student Grant.** *2021 USENIX Annual Technical Conference*, USENIX Association.

2020 **OSDI'20 Student Grant.** *14th USENIX Symposium on Operating Systems Design and Implementation*, USENIX Association.

2020 **Merit Grant.** *Direção-Geral do Ensino Superior (DGES), Portugal.*

2019 **PhD Grant.** *SFRH/BD/146059/2019, Fundação para a Ciência e a Tecnologia, Portugal.*

2018 **Excellence Scholarship.** *University of Minho, Portugal.*

Awards received by advised students

2025 **Pedro Peixoto:** *PhD grant, Fundação para a Ciência e a Tecnologia, Portugal.*

2024 **Diana Rodrigues:** *OERN Masters Award, Ordem dos Engenheiros Região Norte, Portugal.*

Rúben Adão: *PhD grant, Fundação para a Ciência e a Tecnologia, Portugal.*

Talks and Invited Events

Oct 2026 *From Knowledge to Competitiveness: Portugal–US Partnerships 2026 Summit* at *University of Aveiro* (Aveiro, Portugal). Joint summit of the CMU Portugal, MIT Portugal, and UT Austin Portugal Programs; invited as PI of *PATA* (*UT Austin*).

Apr 2026 *Beyond GPU Utilization: Understanding AI Training Energy Consumption* at *InSyDe Workshop, co-located with EuroSys 2026* (Edinburgh, UK)

Jan 2026 *Dagstuhl Seminar 26032: Storage Systems and I/O for Emerging Workloads on HPC Systems* at *Schloss Dagstuhl – Leibniz Center for Informatics* (Wadern, Germany).

Nov 2025 *Redesigning Storage and Operating Systems for Modern Data Centers* at *McGill University, School of Computer Science* (Montreal, Canada)

Feb 2025 *A greener path: Improving AI training energy efficiency through workload-aware GPU tuning* at *CENTRA 8* (Hsinchu, Taiwan)

Jul 2024 *Can software-defined storage solve long-standing HPC storage problems?* at *SIGHPC Computing Continuum* (Virtual)

Nov 2023 *Can software-defined storage solve long-standing HPC storage problems?* at *National Institute of Advanced Industrial Science and Technology* (Tokyo, Japan)

May 2023 *PADLL: Taming Metadata Burstiness of HPC Jobs Through Application-level QoS Control (Poster)* at *ACM EuroSys* (Rome, Italy)

May 2023 *Taming Metadata-intensive HPC Jobs Through Dynamic, Application-agnostic QoS Control* at *IEEE/ACM CCGrid* (Bangalore, India)

Jun 2022 *User-level Software-Defined Storage Data Planes* at *ENSD* (Évora, Portugal)

May 2022 *Building User-level Data Planes with PAIO* at *ADAC Seminar* (Virtual)

May 2022 *User-level Software-Defined Storage Data Planes* at *On the Road to HPC: Major Challenges and New Opportunities* (Webinar, Virtual)

Feb 2022 *PAIO: General, Portable I/O Optimizations With Minor Application Modifications* at *USENIX FAST* (Santa Clara, CA, USA)

Sep 2021 *The Case for Storage Optimization Decoupling in Deep Learning Frameworks* at *REX-IO Workshop* (Virtual)

Oct 2019 *A Case for Dynamically Programmable Storage Background Tasks* at *SRDS Workshops* (Lyon, France)

Sep 2017 *A Practical Framework for Privacy-Preserving NoSQL Databases* at *IEEE SRDS* (Hong Kong)

References

Prof. João Tiago Paulo
INESC TEC & University of Minho
joao.t.paulo@inesctec.pt

Prof. José Orlando Pereira
University of Minho
jop@di.uminho.pt

Dr. Jason Haga
National Institute of Advanced Industrial Science and Technology (AIST)
jh.haga@aist.go.jp

Prof. Oana Balmau
McGill University
oana.balmau@mcgill.ca