PEDRO PALACIOS ALMENDROS

pedro5@illinois.edu ◇ LinkedIn ◇ github.com/palmenros ◇ Google Scholar

EDUCATION

Ph.D. in Computer Science, *University of Illinois at Urbana-Champaign (UIUC)* 2025 - current
Advisor: Josep Torrellas. GPA: 4.0 out of 4.0.

Research topics: computer architecture, ML, distributed systems.
Relevant coursework: Parallel Computer Architecture, Machine Learning Compilers, Cloud Storage Systems, Deep Generative Models.

M.Sc. in Computer Science, *École Polytechnique Fédérale de Lausanne (EPFL)* 2023 - 2025
GPA: 5.96 out of 6. Ranked 1st in Computer Science department and 2nd university-wide.

Relevant coursework: Advanced multiprocessor architecture, Advanced computer architecture, Distributed algorithms, Advanced compiler construction, Systems for data management, VLSI.

B.Sc. in Computer Science (240 ECTS), *Universidad Complutense de Madrid (UCM)* 2018 - 2023
GPA: 9.89 out of 10. Ranked 1st in graduating class. Part of a double degree program.

B.Sc. in Mathematics (240 ECTS), *Universidad Complutense de Madrid (UCM)* 2018 - 2023
GPA: 9.76 out of 10. Ranked 1st in graduating class. Part of a double degree program.

Exchange program during bachelor, *University of California, San Diego (UCSD)* 2021 - 2022
GPA: 4.0 out of 4.0. Awarded provost honors in all three terms.

INDUSTRY EXPERIENCE

IMEC, *Research Intern* January 2025 - June 2025

- Designed and evaluated a near-KV cache Compute-near-Memory accelerator using emerging CCD 3D memory for mobile LLM inference. Patent and publication in progress.

ACADEMIC EXPERIENCE

i-acoma group at UIUC, *Research Assistant* August 2025 - current

- Working with Prof. Josep Torrellas on computer architecture for ML, cloud and storage systems within the SRC JUMP 2.0 ACE Center for Evolvable Computing.

Embedded Systems Laboratory at EPFL, *Research Assistant* June 2024 - January 2025

- Worked on the design, programming, Gem5 simulation and integration into a RISC-V SoC of a low-power mixed-precision ML accelerator supporting structured pruning.

Embedded Systems Laboratory at EPFL, *Research Assistant* March 2024 - June 2024

- Worked in a joint research project with IMEC on micro-architectural simulation, evaluation and optimization of GPUs for 3D stacked transistor technologies.

Dynamically Reconfigurable Hardware Group at UCM, *Research Assistant* 2022 - 2023

- Awarded a Research Collaboration Grant from the Spanish government.
- Developed custom hardware accelerators for real-time hyperspectral image processing on FPGAs.

PUBLICATIONS

- P. Palacios, R. Medina, J.L. Rouas, G. Ansaloni and D. Atienza, *Systolic Arrays and Structured Pruning Co-design for Efficient Transformers in Edge Systems*, in GLSVLSI 2025. Link.
- J.L. Rouas, C. Brazier, L. B. Letaifa, R. Medina, P. Palacios, D. Atienza, G. Ansaloni, *Structured Pruning for Efficient Systolic Array Accelerated Cascade Speech-to-Text Translation* in Interspeech 2025. Link.
- P. Palacios, D. Báscones, C. González and D. Mozos, *A Real-Time FPGA Implementation of the LCMV Algorithm for Target Classification in Hyperspectral Images Using LDL Decomposition*, in IEEE Transactions on Geoscience and Remote Sensing, 2024. Link.

PATENTS

- P. Palacios, B. Wang, J. Klein, F. Garcia, G. Ansaloni, D. Atienza, *Transformer Auto-Regressive Attention Accelerator Using 3D String-Memory Key-Value Cache*, Patent in progress, 2025.

WORKSHOP PAPERS

- C. Block, P. Palacios, A. Farrell, G. Gerogiannis, J. Torrellas, *Performance-Driven Composite Prefetching with Bandits* in DPC4 at HPCA 2026.
- P. Palacios, J. Klein, A. Sharma, D. Abdi, G. Ansaloni, F. Garcia, D. Biswas, D. Atienza, J. Myers, *Full Stack Framework for Architectural Exploration of IGZO-Based Compute-near-DRAM* in IMEC PTW 2025.
- P. Palacios, R. Medina, G. Ansaloni, D. Atienza, *HEEPstor: an Open-Hardware Co-design Framework for Quantized Machine Learning at the Edge* in Open-Source Hardware Workshop at ACM CF 2025.

PRIZES AND AWARDS

- Awarded the *"Andrew & Shana Laursen Fellowship"* for top incoming UIUC grad students ($30,000) 2025
- Awarded the *"2nd Best Master Academic Record"* university-wide across all departments by EPFL 2025
- Awarded the extremely competitive *"La Caixa" Postgraduate Abroad Fellowship* ($65,000) 2023-2025
- Received the *Award for the Best Academic Record* by Universidad Complutense de Madrid 2023
- Awarded the *Excellence Scholarship* four times by the Community of Madrid ($10,000) 2019-2023
- Awarded the competitive *UC Exchange Scholarship* by UCM to fund study abroad at UCSD 2021-2022
- Awarded a travel grant to attend ISCA 2024 2024
- Participated in and achieved various positions in several competitive programming events, and hardware competitions, as well as national mathematics, computer science, and physics olympiads 2016-2023

OTHER PROJECTS

- **Master's Thesis**. *System-Level Exploration of Emerging Memory Technologies for Compute-near-Memory Architectures.* Developed a modeling framework for emerging IGZO eDRAM and CCD 3D memories, spanning from device-level simulations to full system architectural simulation. Designed a custom CnM architecture to map LLM inference workloads to emerging 3D CCD memory.
- **Mathematics Bachelor's Thesis (awarded best thesis).** *Algebraic methods for arithmetic circuit analysis in zero-knowledge proofs.* Created an algorithm based on Gröbner bases to formally verify the safety of cryptographic zero-knowledge proof programs and implemented it on top of the open-source compiler Circom. Link.
- **Asgard**: Research project for Prof. Babak Falsafi's Advanced Multiprocessor Architecture course at EPFL. Top ranked student in class. Hardware-partitioning technique for terabyte-scale bare-metal cloud servers using the Midgard address space that improves granularity for core assignment in bare-metal cloud and reduces Last-Level Cache communication and physical sharing, maintaining high associativity. Link.
- **RISC-V SoC Implementation in 28nm FDSOI**: Worked on a microelectronic implementation of a RISC-V processor using the FDSOI 28nm process with the Department of Computer Architecture at UCM. Integrated the open-source SweRV RISC-V core into an SoC, and developed TCL scripts for synthesis and physical design.
- **Decentralized Rich-Text Encyclopedia**: Implemented a scalable, decentralized P2P encyclopedia in Go, with Byzantine fault-tolerant vote propagation, distributed reputation management using EigenTrust, and a custom rich-text CRDT for concurrent article modifications. Censorship-resistant using a Web of Trust. Link.

TOOLS AND SKILLS

C++, Rust, Java, Scala, GoLang, CUDA, Python, Bash, VHDL, SystemVerilog, Chisel, RISC-V, Gem5, AccelSim, Ramulator, DRAMPower, QEMU, Vitis HLS, Vivado, Cadence Genus, Cadence Innovus, perf, Valgrind, Linux Driver Development, LLVM, MLIR, XLA, Triton, Amazon Trainium NKI, AMD MLIR-AIE, PyTorch, JAX, Spark, Memcached, Docker, Azure, AWS.