SORS: Automated cross-architecture CARM modeling and application profiling

Abstract

In recent years, HPC systems have become increasingly complex and heterogeneous, making application development and optimization challenging. To this respect, intuitive performance models like the Cache-aware Roofline Model (CARM) offer effective guidance by providing insights into bottlenecks that limit the application's ability to reach the system's maximum performance. The current landscape of CARM-enabled tools covers either vendor-specific (Intel Advisor), not sufficiently developed (AMD) or simply non-existing (ARM, RISC-V) tools.

The focus of this talk is the CARM Tool, which was developed to address this problem, by extending CARM support to all major CPU architectures and ISAs, i.e., x86 (Intel, AMD), ARM, and RISC-V. The proposed tool includes automatically generated assembly microbenchmarks, specifically tailored to cover a full performance spectrum of modern CPUs (from scalar to all supported vector ISA extensions) for both computational units and all memory hierarchy levels. The tool also provides application profiling capabilities in the scope of the CARM, to facilitate application optimizations using the CARM's insight.

Short Bio

José Morgado is an Assistant Researcher at INESC-ID, Lisbon, specialized in High Performance Computing (HPC) systems. José earned his Master's degree in High Performance Computing Systems from Instituto Superior Técnico, having focused on computer architectures and microbenchmarking during his studies. His Master's thesis led to the creation of a microbenchmarking tool that accurately generates the Cache-Aware Roofline Model (CARM) for various CPU architectures and vendors, based on automatically generated assembly microbenchmarks, which also resulted in a scientific publication at the IISWC24 conference. He has also enhanced this tool with a graphical user interface and enabled detailed performance analysis through the use of performance counters and dynamic binary instrumentation. Shortly after completing his thesis, José contributed to the SparCity project, where he developed the live-CARM feature within one of the project's frameworks, this work was subsequently published in a scientific paper, which he co-authored. Currently, José is engaged in actively contributing to several European projects and institutions (POP3, SYCLOPS, VI-HPS, CERN, BSC, Eurecom), including collaborations with the Barcelona Supercomputing Center, where he is working on integrating support for CARM-based application analysis in the Paraver ecosystem.

Speakers

Speaker: José Morgado. Assistant Researcher at INESC-ID, Lisbon, specialized in High Performance Computing (HPC) systems
Host: Petar Radojkovic. Established Researcher. Computer Sciences - Memory technologies, BSC Barcelona Supercomputing Center - Centro Nacional de Supercomputación