From Memory Partitioning to Management through Fine-grained Profiling and Control

Renato Mancuso
Boston University

Time: March 4th, 1pm-2:15pm,

Over the last decade, embedded computing platforms have exploded in complexity. The push for complexity has been driven by the need for context-awareness in next-generation cyber-physical systems. That is, the ability to exploit knowledge of the environment and to make complex decisions based on a multitude of sensory streams. Unfortunately, as platforms grow in complexity to improve context-awareness, the inter-play between concurrent software components and the underlying hardware becomes hard to predict and to reason about. The latter can be thought of as the capacity to achieve self-awareness. Therefore, there exists a fundamental tension between context- and self-awareness. The lack of temporal isolation in modern platforms has shaken the foundations of real-time theory, embedded systems design, verification, and validation. Seminal results have been accomplished to mitigate temporal interference and achieve strong performance isolation via hardware resource partitioning. But the problem largely remains an open research question. At its core, the issue of temporal interference shares many similarities with a class of problems in security threat identification and mitigation — namely time-based side-channel attacks. Unsurprisingly, both challenges trace back to a lack of self-awareness in modern platforms.

This begs the question: can we set aside resource partitioning as “poor man’s management” and elevate self-awareness in modern embedded systems instead? In this talk, I will walk you through some of the milestones in hardware resource partitioning that have led to important changes in the way we design modern operating systems and hypervisors. I will then review the latest advancements propelled by my research lab in techniques to build knowledge of the application workload, which is a crucial stepping-stone for self-awareness. I will also review the fundamental mechanisms that can allow exerting fine-grained monitoring and control over hardware resources and data-flows. Finally, I will discuss an overarching vision for how self-awareness can be achieved in current and future embedded platforms.

Short Bio: Renato Mancuso is an assistant professor in the Computer Science Department at Boston University (BU). Before joining BU in 2017, Renato received his Ph.D. from the Department of Computer Science at the University of Illinois at Urbana-Champaign. He is interested in high-performance cyber-physical systems, with a specific focus on techniques to enforce strong performance isolation and temporal predictability in multi-core, accelerator-enabled systems. He has published more than 40 papers in major conferences and journals. His papers received multiple awards in top conferences in the field of real-time and embedded systems. Renato also is the information director of the ACM SIGBED special interest group on embedded systems. His research is supported by the NSF, Bosch GmbH, and Red Hat. He is a member of the IEEE.