



DISSERTATION DEFENSE



ARUN SUBRAMANIYAN

Accelerated Systems for Pattern Matching

Friday, October 22, 2021

9:30–11:30 AM

3901 Beyster, [Virtual](#)

(Passcode: 316621)

Meeting ID: 970 4047 0726

ABSTRACT: Pattern matching forms the core of many applications and contributes to a large fraction of their execution time. For instance, scanning the human reference genome for motifs to identify variations between individuals can take several days on a multi-core processor and parsing activities in the front-end of web browsers can account for a significant fraction of web page loading time. While general-purpose processors have been optimized for regular, data-parallel workloads, they are not well suited for pattern matching. These workloads often perform irregular memory accesses and spend significant amounts of time and energy in data movement from storage to compute units and are bottlenecked by memory bandwidth.

In this talk, I will demonstrate how hardware-software co-design can be used to alleviate some of the computational challenges presented by these pattern matching workloads. First, I will present a new hardware design that allows embarrassingly sequential finite state automata, a common computation model used for pattern matching, to be executed in parallel in a DRAM-based in-memory accelerator. Next, I will present the Cache Automaton architecture, which repurposes CPU last-level caches for automata processing. The second part of the talk takes a deep dive into genomics analysis, an application domain which includes pattern matching as a key computational component. Here, I will present our approach to accelerate string matching in a time-consuming genomics analysis workload - read alignment, which matches each of the billions of short fragments of DNA emitted by the sequencer (called reads) against a reference genome. Finally, I will present our efforts to characterize the pattern matching landscape in genomics and highlight opportunities for the development of new domain specific architectures customized for genomics analysis.

CHAIR: Prof. Reetuparna Das