

CRITICAL ANALYSIS AND EVALUATION OF DIFFERENT AUTOMATA PROCESSING ACCELERATORS ON LARGE-SCALE DATASETS

Marziyeh Nourian

Dr. Michela Becchi, Thesis Supervisor

ABSTRACT

Many established and emerging applications perform at their core some form of pattern matching, a computation that maps naturally onto finite automata abstractions. As a consequence, in recent years there has been a substantial amount of work on high-speed automata processing, which has led to a number of implementations targeting a variety of parallel platforms: from multicore CPUs, to GPUs, to FPGAs, to ASICs, to Network Processors. More recently, Micron has announced its Automata Processor (AP), a DRAM-based accelerator of non-deterministic finite automata (NFAs). Despite the abundance of work in this domain, the advantages and disadvantages of different automata processing accelerators and the innovation space in this area are still unclear.

In this work we target this problem. In particular, to allow an apples-to-apples comparison, we focus on NFA acceleration on three platforms: GPUs, FPGAs and Micron's AP. We discuss the automata optimizations that are applicable to all three platforms. We perform an evaluation on large-scale datasets: to this end, we propose an NFA partitioning algorithm that minimizes the number of state replications required to maintain functional equivalence with an unpartitioned NFA, and we evaluate the scalability of each implementation to both large NFAs and large numbers of input streams. Our experimental evaluation covers resource utilization, throughput, and preprocessing cost.