Analytic Query Processing using Associative Computing

Abstract:
Database analytic query workloads are heavy consumers of data-center cycles, and there is constant demand to improve their performance. Associative processors (AP) have re-emerged as an attractive architecture that offers very large data-level parallelism that can be used to implement a wide range of general-purpose operations. Associative processing is based primarily on efficient search and bulk update operations. Analytic query workloads benefit from data parallel execution and often feature both search and bulk update operations. In this paper, we investigate how amenable APs are to improving the performance of analytic query workloads. For this study, we use the recently proposed Content-Addressable Processing Engine (CAPE) framework. CAPE is an AP core that is highly programmable via the RISC-V ISA with standard vector extensions. By mapping key database operators to CAPE and introducing AP-aware changes to the query optimizer, we show that CAPE is a good match for database analytic workloads. We also propose a set of database-aware microarchitectural changes to CAPE to further improve performance. Overall, CAPE achieves a 10.8× speedup on average (up to 61.1×) on the SSB benchmark (a suite of 13 queries) compared to an iso-area aggressive out-of-order processor with AVX-512 SIMD support.

Bio:
Yannis recently received his Ph.D. from the Computer Sciences Department at the University of Wisconsin-Madison. He was advised by Prof. Jignesh Patel and his research focused on making data processing efficient for modern and emerging hardware. Prior to his Ph.D., he studied at the University of Athens in Greece. Yannis will join the System Research Group at Google in the fall.

Date and time: Tuesday, October 04th, 2022, 11.00 am
Location: Pérolles 21, room D230, Bd de Pérolles 90, Fribourg
Contact person: Prof. Philippe Cudré-Mauroux

The colloquium is free and open to the public.