SORS: Combining Load Operations

Objectives

Abstract: General-purpose processors continue to proliferate, and microarchitectural innovations are going to be critical to continue improving their performance and energy efficiency. In this talk we will present a novel technique for combining multiple load instructions in order to optimize their execution. A simple technique is used to learn and predict the number of contemporaneous accesses to a region of memory and classify a particular dynamic load into a normal or a fat load. Fat loads bring in additional data into Contemporaneous Load Access Registers (CLARs), from where other contemporaneous loads could be serviced without accessing the L1 cache. Going further, many such loads (50-60%) could be processed in the register renaming stage. This results in a significant reduction in the number of L1 cache accesses, with commensurate energy savings. In several cases the reduced latency for such loads results in an earlier resolution of some mispredicted branches, and a reduction in the number of wrong-path instructions, especially loads.

Short Bio: Guri Sohi has been a

faculty member at the University of Wisconsin-Madison since 1985 where he currently a John P. Morgridge Professor and a Vilas Research Professor. His research has been in the design of high-performance microprocessors and computer systems. Results from his research can be found in almost every high-end microprocessor in the market today.

Speakers

Speaker: Professor Guri Sohi, University of Wisconsin-Madison
Host: Osman Unsal, Computer Architecture For Parallel Paradigms Group Manager, CS, BSC

Barcelona Supercomputing Center - Centro Nacional de Supercomputación

---