

MIT9904-04

Malleable Architectures for Adaptive Computing

Arvind, Larry Rudolph and Srinivas Devadas,

Project Overview

Neither general-purpose microprocessors nor digital signal processors meet all the needs of intelligent personal devices, multimedia players and recorders, and advanced communication applications. Designing special purpose chips for each application is too expensive to be a feasible solution. We believe that the best approach is to make current general-purpose processors more malleable.

During the past year, we have been investigating ways of making better use of on-chip cache memory. This research has suggested that augmenting the L2 cache with processing power can dramatically improve general-purpose processor performance for the new breed of stream-based applications. This can be done by modifying a general-purpose microprocessor making it more malleable and optimized for a wide range of applications. In particular, we propose to (1) augment the ISA with vector operations, fine-grained memory fence operations, and cache management instructions, (2) simplify the processor core, and (3) associate multiple functional units with the banks of L2 cache memory.

In conjunction with the malleable cache research, we have also been investigating a revolutionary technology for designing hardware and firmware from high-level specifications. The approach is to synthesize "malleable" processors from traditional ones. The processor will contain most of the functional units, data paths, and cache memory found in current general-purpose microprocessors, but they will be surrounded with pieces of reconfigurable logic. The instruction sets of the processors are tailored to each application so as to significantly improve either the performance or the power dissipated during the execution of the application. Synthesis of instruction sets is made possible by the development of an architecture exploration system for programmable processors. This technology can dramatically reduce the time to market in sectors where the standards are changing too quickly or where functionality evolution is too rapid for traditional hardware design. Further, processor caches can be reconfigured in a dynamic

manner so as to improve hit rates for multimedia streaming data. This reconfiguration is made possible by implementing several hardware mechanisms such as column and curious caching into the processor cache. Column caching provides control as to where items are stored in a cache and curious caching allows the cache to fetch data that has been placed on the bus.

During the coming year, we hope to investigate such a “smart-cache” architecture that will support vector type operations directly in the L2 cache and the scalar operations in a simplified, but traditional processor core. This malleable processor will execute traditional programs, but in a significantly more efficient and higher performance manner. The architecture will be specified using our new specification technology, and demonstrate its performance on a set of stream-based application programs.

Hiroshi Sawada of NTT will be working at LCS to help with the architectural specification and investigation of its usefulness for three major applications: compression applications (JPEG and MPEG), the speech analysis software developed in Victor Zue’s Spoken Language Systems group at the MIT LCS lab, and Image understanding in Paul Viola’s Learning and Vision Group at the MIT AI lab.