Abstract

Active Messages: an Efficient Communication Architecture for Multiprocessors

by

Thorsten von Eicken

Doctor of Philosophy in Computer Science
University of California at Berkeley
Professor David E. Culler, Co-Chair
Professor John Wawrzynek, Co-Chair

This dissertation presents Active Messages, a new communication architecture for massively parallel multiprocessors. Active Messages achieves an order of magnitude performance improvement over conventional communication layers and efficiently supports a variety of parallel programming models, including message passing, shared memory and dataflow. The key to this success is that Active Messages is predicated on the use of high-level parallel languages and takes the entire system into account, from the micro-architectural level up to the language system. This dissertation develops a conceptual framework around the notion of a communication architecture extending the sequential instruction set architecture. The focus is on integrating communication into the compilation process and allowing trade-offs to be made across all hardware and software layers of the system. Understanding how the various layers of abstraction interact allows the key communications issues to be addressed at the right level in an approach analogous to RISC architectures.

The analysis of traditional systems (e.g., message passing, shared memory, message driven, and dataflow) concludes that these provide too complex and rigid primitives that are tailored to a specific programming model. In contrast, Active Messages provides simple communication primitives appropriate for code generation and decouples storage allocation and sophisticated scheduling from communication proper. This enables powerful compiler optimizations and results in a more efficient system as demonstrated through the analysis of programs written in Split-C and Id90. On a single platform, the CM-5, Active Messages supports message passing, message driven, dataflow, and NUMA shared memory programming models today as efficiently as more specialized hardware, and, coupled with hardware support for a global address space, is appropriate for implementing cache-coherent shared memory.

The simplicity of Active Messages make it an attractive basis for the development of new hardware support for communication. This dissertation analyzes the timing of Active Messages in detail on the CM-5 and the nCUBE/2 to uncover simple hardware changes which could double the communication performance. A number of ongoing multiprocessor hardware developments use Active Messages and promise more dramatic improvements. To date, the Active Messages communication architecture offers the most versatile and efficient avenue towards multiprocessors which support a full spectrum of parallel programming languages.