Communication and Memory Requirements as the Basis for Mapping Task and Data Parallel Programs

Jaspal Subhlok David R. O'Hallaron Thomas Gross
Peter A. Dinda Jon Webb

January 1994
CMU-CS-94-106

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

For a wide variety of applications, both task and data parallelism must be exploited to achieve the best possible performance on a multicomputer. Recent research has underlined the importance of exploiting task and data parallelism in a single compiler framework, and such a compiler can map a single source program in many different ways onto a parallel machine. There are several complex tradeoffs between task and data parallelism, depending on the characteristics of the program to be executed, most significantly the memory and communication requirements, and the performance parameters of the target parallel machine. In this paper, we isolate and examine the specific characteristics of programs that determine the performance for different mappings on a parallel machine, and present a framework for obtaining a good mapping. The framework is applicable to applications that process a stream of input, and whose computation structure is fairly static and predictable. We describe three applications that were developed with our compiler: fast Fourier transforms, narrowband tracking radar and multibaseline stereo, examine the tradeoffs between various mappings for them, and show how the framework was used to obtain efficient mappings. The automation of this framework is described in related publications.

This research was sponsored by the Advanced Research Projects Agency/CSTO monitored by SPAWAR under contract N00039-93-C-0152, and also by the Air Force Office of Scientific Research under contract F49620–92–J–0131.

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either express or impied, of the U.S. government.