Continuous technology miniaturization allows to build massively parallel embedded computer architectures within a single silicon chip. Programming that leverages the abundant parallelism in such architectures, however, is very difficult, tedious, and error-prone. Thus, compiler support is paramount. We therefore present LoopInvader, a loop compiler for a particular class of massively parallel processor arrays: Tightly coupled processor arrays (TCPAs) [2].

TCPAs consist of a two-dimensional array of VLIW processing elements (PEs) and several peripheral components that enable zero-overhead loops. In particular, a global controller (GC) generates synchronized control signals that govern the control flow of the PEs, removing control overhead from the loops; address generators (AG) produce the necessary addresses for feeding the PEs with data from reconfigurable buffers, removing addressing overhead. Moreover, the PEs are connected to their neighbors via a circuit-switched interconnection network that is reconfigurable at runtime to optimally accommodate the running application.

Figure 1 depicts an overview of our high-level programming methodology. We describe programs in a domain-specific functional language called PAULA that is based on dynamic piecewise linear/regular algorithms (DPLA) [3], a mathematical representation of loop programs. For the parallelization and mapping of such algorithms onto TCPAs we use symbolic partitioning techniques [5] in the polyhedral model: Instead of using fixed tile sizes, our symbolic partitioning technique is able to keep the size of the input data and the number of PEs symbolic until runtime. This provides applications more flexibility and is important in resource-aware computing paradigms such as invasive computing [4]. Other approaches are both time-consuming (e.g., dynamic recompilation) and costly (e.g., pre-compiling multiple variants) on embedded systems.

After mapping, the compiler generates a configuration stream comprising assembly code for the PEs, interconnect configuration, address generator configuration and global controller configuration. Because the PEs offer only small instruction memories, we developed an approach to generate code that is independent of the problem size [1]. This is achieved by finding processors and program blocks within processors that share the same code and appropriately combining it into loops.

As the PEs are interconnected by a circuit-switched interconnection network, the compiler also generates all necessary configuration information. For preserving a given schedule of instructions, code for the GC is generated such that the repetitive execution of each unique program block does not cause any extra cycles.

This orchestration enables the execution of nested loop programs with zero-overhead loop, not only for innermost loops but also for all static conditions in arbitrary multi-dimensional data flow. Finally, the generated configuration stream can be loaded either onto an actual TCPA synthesized on an FPGA or into a cycle-accurate simulator that we developed in tandem with the architecture and compiler for testing and verification.

ACKNOWLEDGMENT

This work was supported by the German Research Foundation (DFG) as part of the Transregional Collaborative Research Center “Invasive Computing” (SFB/TR 89).

REFERENCES


