“Screwdriver” Cache Coherence for Multicore Processors

Peter Hornyack
Faculty Advisor: Dr. Milo M. K. Marlin

Abstract

- “Screwdriver” is a cache coherence protocol that avoids race conditions and emphasizes scalability while being implementable on an FPGA.

Multicore Processors

- Link together multiple cores that work in parallel on the same chip.
- Intel and AMD produce dual-core and quad-core processors for consumers today.

Shared Memory

- As each processor reads from and writes to memory:
  - Fast, local cache memories bypass the slow, shared main memory.
  - Cache coherence protocols ensure that the data all processors use is valid.

The Screwdriver protocol

- Uses a single central memory controller that:
  - Handles memory requests.
  - Issues coherence messages.
  - Limits race conditions.
- Emphasizes scalability.
  - The ability of a protocol to support an increasing number of processor cores.
- Inspired by AMD Hammer protocol, which uses one memory controller per processor.

FPGAs

- Hardware description language (HDL) used to connect logic components into a working multicore design.
  - AND, OR, NOT gates
  - Registers
  - Memory blocks
  - ALUs
- HDL design is synthesized and downloaded to FPGA chip (field-programmable gate array).
- Programming is temporary and process can be repeated.

Analysis of Screwdriver protocol

- Scalability:
  - No theoretical limit to number of processor cores connected to memory controller.
- Multicore prototypes analyzed:
  - Simulation: 16 core design.
  - FPGA implementation: 8 core design.
- “Relative work”:
  - Amount of work done relative to a single baseline processor.

Conclusions

- Increasing number of processors demonstrates Screwdriver’s limited bandwidth:
  - Each processor waits longer to become the active processor.
  - Greater number of processors -> greater number of invalidation requests -> more cache misses.
  - Effects greater for memory-intensive benchmark.
- Screwdriver succeeds in being easily scalable:
  - Successful FPGA implementation of 8 core design.
  - Successful simulation of 16 core design.
  - Larger designs limited only by space on FPGA chip and simulation time required.
- Single memory controller helps avoid race conditions, but at the expense of memory request bandwidth.