Computer Systems Organization

CSE 521/560M
Lecture 14
Prof. Patrick Crowley

Plan for Today

- Questions
- Today’s discussion
Predicated Instructions

- Removal of conditional branches via the conditional execution (or completion) of instructions based on the result of a Boolean condition.
  - Has its roots in vector machines (CRAY-1 had a vector mask which held the sign of each vector element to accelerate abs value operation)
  - Implementations often restrict predication to conditional moves (e.g., Alpha)
  - The PA-RISC allowed some instructions to nullify the next one
  - More generally, a guarded execution model uses a special instruction to conditionally control subsequent ones
  - The Cydra 5 VLIW machine allowed every instruction to be conditional

Conditional Moves

- CMOVZ R2, R3, R1
  - Move R3 to R2 if R1 = 0

- Compiler example:
  - If (A==0) { S=T; }  
  - Read as:
    - If (R1==0) { R2=R3; }
  - Implement with:
    - CMOVZ R2, R3, R1
Multithreading Primer

- Idea: Exploit parallelism available between threads
- Each thread needs HW support
  - Register File
  - PC, other special registers
  - Memory system bookkeeping (e.g., page table)
  - Fast context switches

Thread Granularity

- Fine-grained
  - Context switches every cycle, round-robin schedule that skips stalled threads
  - Pro: Hides short and long latencies
  - Con: Each thread issues a max of 1 instruction per issue “rotation” (i.e., many threads can provide good throughput but poor individual latency)

- Coarse-grained
  - Context switches only on long-latency stalls (e.g., cache misses)
  - Pro: Individual threads can have better latency
  - Con: Cannot hide short latencies due to pipeline flushes
Simultaneous Multithreading

- Idea: mix fine-grain multithreading with a wide-issue speculative superscalar
  - HW can already handle dependencies, just add thread id bits
  - Use multiple issue from multiple threads since single threads rarely have enough ILP to fill issue width

Visual Comparison

From: Converting Thread-Level Parallelism to Instruction-Level Parallelism via Simultaneous Multithreading, LO et al.
Assignment

• Readings
  – For Wednesday
    • H&P: 6.9
    • H&P: 5.1-5.2
    • Ash: Ch. 10
  – For Monday
    • Commentary: The superblock: an effective technique for VLIW and superscalar compilation