Exploiting More ILP

- ILP = ____________________________ (parallelism within a single program)
- How can we exploit more ILP?
  1. ________________________________ (Split execution into many stages)
  2. ________________________________ (Start executing more than one instruction each cycle)

Multiple Issue Processors

- Key metric: CPI \rightarrow IPC
- Key questions:
  1. What set of instructions can be issued together?
  2. Who decides which instructions to issue together?
    - Static multiple issue
    - Dynamic multiple issue

Ideas for improving Multiple Issue

1. Non-blocking caches
2. Speculation
3. Loop unrolling
Example: Loop unrolling

Loop:  
lw  $t0, 0($s1)  
sw  $t0, 0($s2)  
addi $s1, $s1, -4  
addi $s2, $s2, -4  
bne $s1, $zero, Loop

Loop:  
lw  $t0, 0($s1)  
lw  $t1, 4($s1)  
lw  $t2, 8($s1)  
lw  $t3, 12($s1)  
sw  $t0, 0($s2)  
sw  $t1, 4($s2)  
sw  $t2, 8($s2)  
sw  $t3, 12($s2)  
addi $s1, $s1, -16  
addi $s2, $s2, -16  
bne $s1, $zero, Loop

Why is this a good idea?

Chapter 6 Summary

- Slow  Faster
Instructions per clock (IPC = 1/CPI)

- Deeply pipelined
  Multiple issue with deep pipeline
  (Section 6.10)

- Single-cycle
  (Section 5.4)

- Multicycle
  Pipelined
  Multiple-issue pipelined
  (Section 6.9)

Multiprocessors (Chapter 9)

- Idea: create powerful computers by connecting many smaller ones
  
good news: works for timesharing (better than supercomputer)
  
bad news: its really hard to write good concurrent programs
  
many commercial failures

Who? When? Why?

- “For over a decade prophets have voiced the contention that the organization of a single computer has reached its limits and that truly significant advances can be made only by interconnection of a multiplicity of computers in such a manner as to permit cooperative solution…. Demonstration is made of the continued validity of the single processor approach…”

- “…it appears that the long-term direction will be to use increased silicon to build multiple processors on a single chip.”
Flynn’s Taxonomy (1966)

1. Single instruction stream, single data stream
2. Single instruction stream, multiple data streams
3. Multiple instruction streams, single data stream
4. Multiple instruction streams, multiple data streams

Question #1: How do parallel processors share data?

1. Shared variables in memory
2. Send explicit messages between processors

Clusters

- Constructed from whole computers
- Independent, scalable networks
- Strengths:
  - Many applications amenable to loosely coupled machines
  - Exploit local area networks
  - Cost effective / Easy to expand
- Weaknesses:
  - Administration costs not necessarily lower
  - Connected using I/O bus
- Highly available due to separation of memories
- Approach taken by Google etc.
A Whirlwind tour of Chip Multiprocessors and Multithreading

Slides from Joel Emer’s talk at Microprocessor Forum

Instruction Issue

Reduced function unit utilization due to...

Superscalar Issue

Superscalar leads to more performance, but lower utilization

Chip Multiprocessor

Limited utilization when only running one thread
Concluding Remarks

- Evolution vs. Revolution

  “More often the expense of innovation comes from being too disruptive to computer users”

  “Acceptance of hardware ideas requires acceptance by software people; therefore hardware people should learn about software. And if software people want good machines, they must learn more about hardware to be able to communicate with and thereby influence hardware engineers.”