ECSE 425 Lecture 24: The Limits of ILP

H&P Chapter 3

Last Time

- Virtual Memory
  - Why do we need it?
  - How do we meet those needs?
  - What do we do to make it fast?

Today

- Studies of the Limitations of ILP
  - Chapter 3.1-3.3

Recall ILP techniques

- Goal: CPI < 1, preserving the programming model
  - Pipelining
  - Loop unrolling
  - Branch prediction (static and dynamic)
  - Dynamic scheduling
  - Speculation
  - Multiple issue
- Modern problems:
  - Bigger processor, faster clock, but same basic structure
  - Complex designs, that are power-hungry and hot
  - Growing gap between peak and delivered performance

Overcoming the Limits of Available ILP

- Can these challenges be overcome?
- What if there are
  - Substantial advances in compiler technology, and
  - Significantly new and different hardware techniques?
- Studies have shown that
  - Even in the best case, ILP is limited, and that
  - Realistic hardware is unlikely to overcome these limits in the near future.
- We'll evaluate ILP perfect hardware, and observe what happens as we make it more realistic

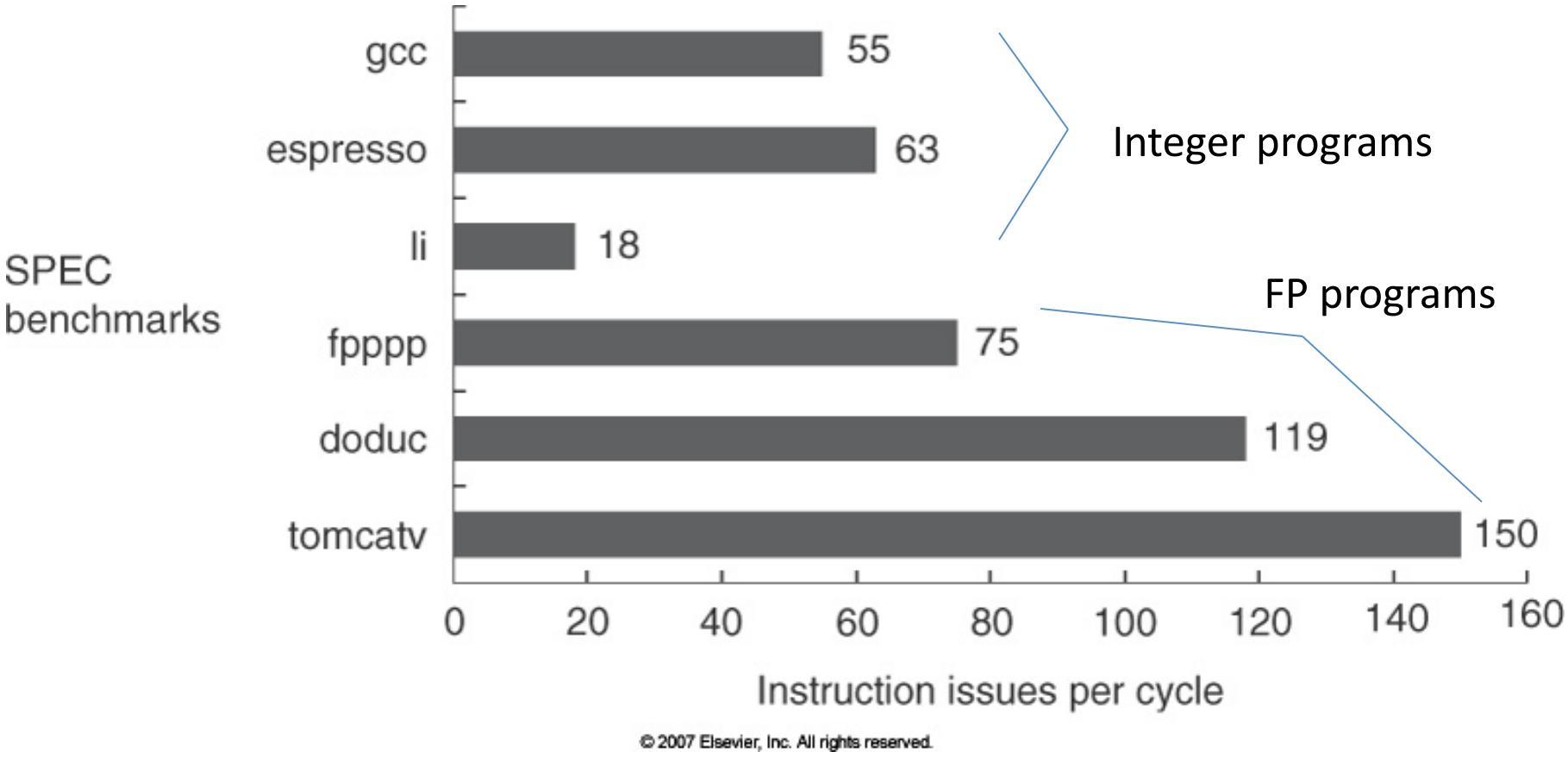
An Ideal Hardware model

1. *Register renaming* – infinite virtual registers
   - all register WAW & WAR hazards are avoided
2. *Branch prediction* – perfect; no mispredictions
3. *Jump prediction* – all jumps perfectly predicted (returns, case statements). (2) and (3) mean
   - no control dependencies
   - perfect speculation, unbounded buffer of instructions
4. *Memory-address alias analysis* – all addresses are known, and all accesses to different addresses can be re-ordered
   - (1) and (4) mean all hazards are eliminated but RAW
5. *Perfect caches* – 1 cycle latency for all instructions (FP*,/); unlimited instructions issued/clock cycle

Model Comparison

	Ideal Model	**IBM Power 5**
Instructions Issued per clock	Infinite	4 can be issued; 6 can begin execution
Instruction Window Size (in flight)	Infinite	200, inc. up to 32 loads or stores
Renaming Registers	Infinite	88 Integer + 88 Floating Point (In addition to 64 arch regs)
Branch Prediction	Perfect	2% to 6% misprediction (Tournament Branch Predictor)
Cache	Perfect	64KI, 32KD, 1.92MB L2, 36 MB L3
Memory Alias Analysis	Perfect	??

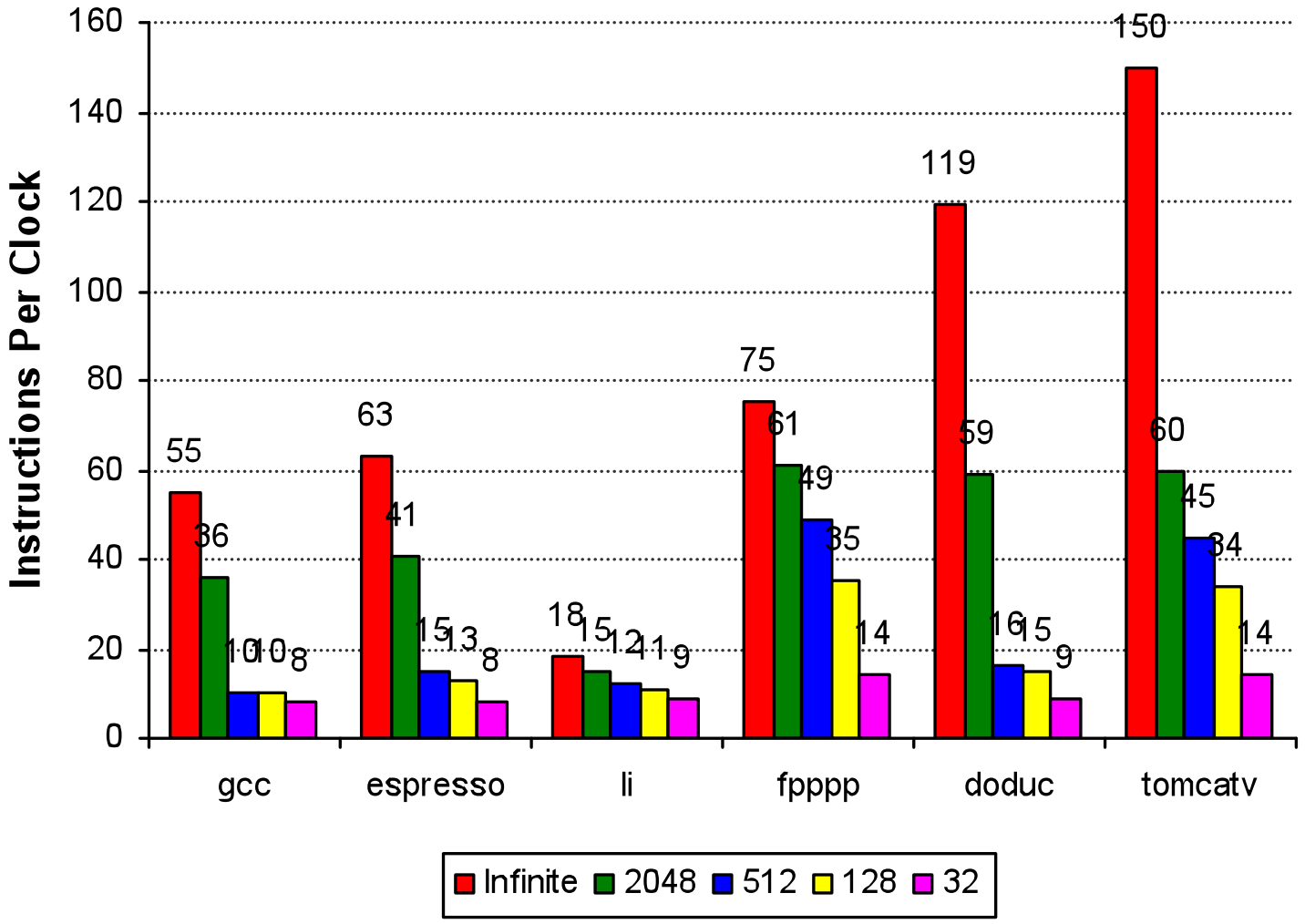
Upper Limit of ILP: Ideal Machine



Realistic Window Size

	New Model	Ideal Model	Power 5
Instructions Issued per CC	Infinite	Infinite	4 can be issued; 6 can begin execution
Instruction Window Size	2K, 512, 128, 32	Infinite	200, inc. up to 32 loads or stores
Renaming Registers	Infinite	Infinite	88 Integer + 88 FP (In addition to 64 arch regs)
Branch Prediction	Perfect	Perfect	2% to 6% misprediction (Tournament Branch Predictor)
Cache	Perfect	Perfect	64KI, 32KD, 1.92MB L2, 36 MB L3
Memory Alias	Perfect	Perfect	??

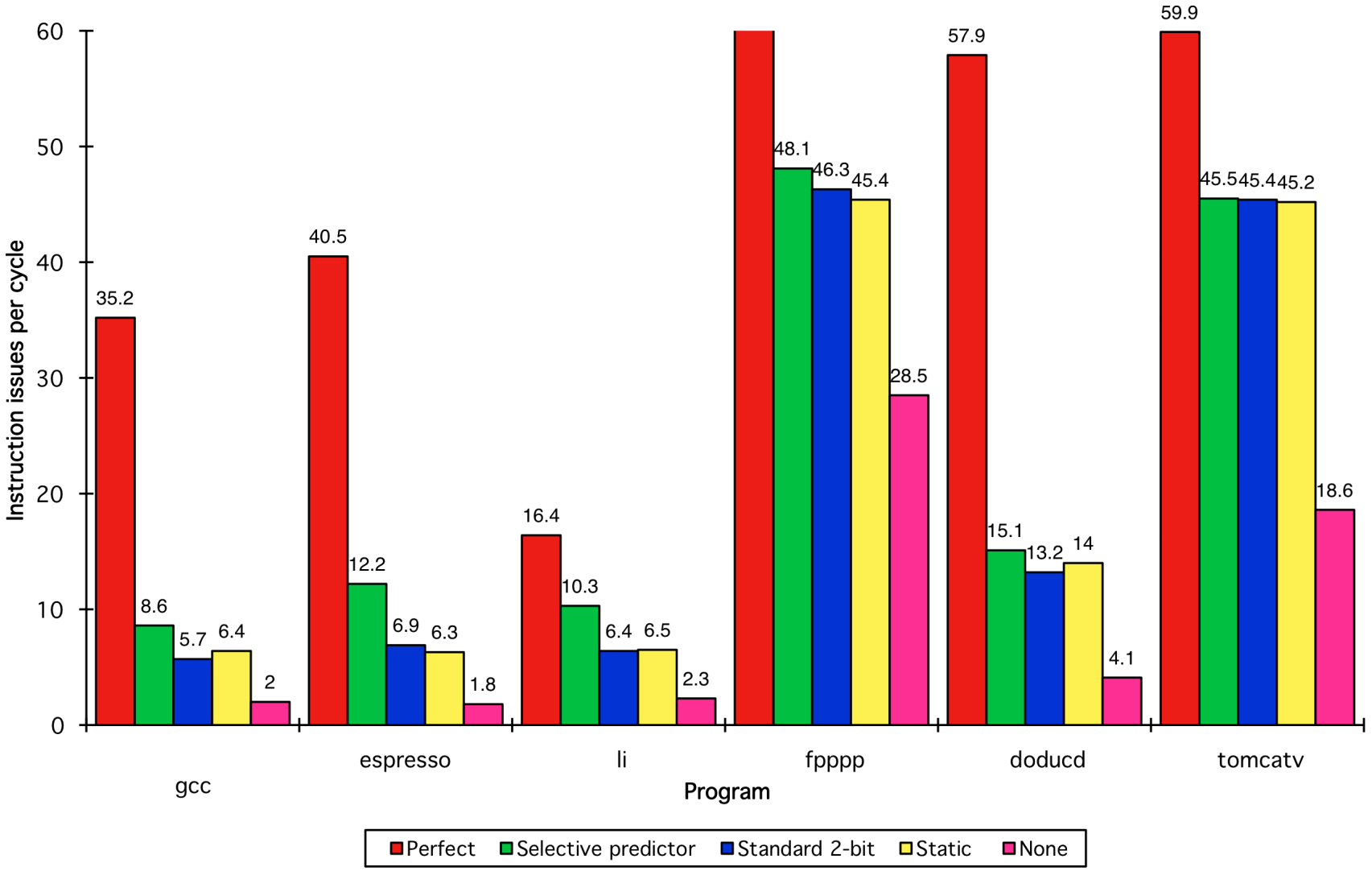
Realistic Window Size, Results



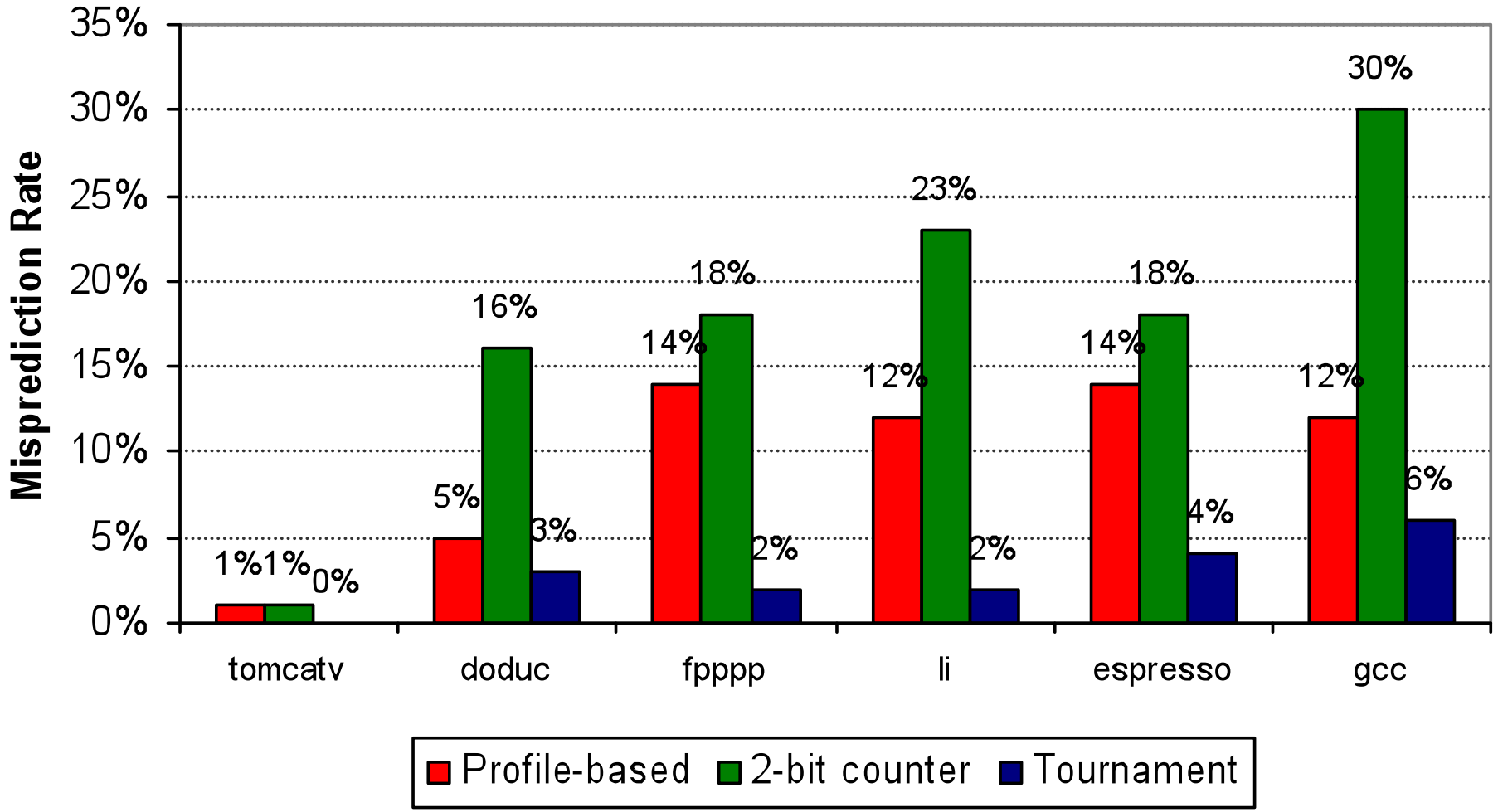
Realistic Branch Prediction

	New Model	Ideal Model	Power 5
Instructions Issued per CC	64	Infinite	4 can be issued; 6 can begin execution
Instruction Window Size	2048	Infinite	200, inc. up to 32 loads or stores
Renaming Registers	Infinite	Infinite	88 Integer + 88 FP (In addition to 64 arch regs)
Branch Prediction	8K Tournament, 512 2-bit, profiling, none	Perfect	2% to 6% misprediction (Tournament Branch Predictor)
Cache	Perfect	Perfect	64KI, 32KD, 1.92MB L2, 36 MB L3
Memory Alias	Perfect	Perfect	??

Realistic Branch Prediction, Results



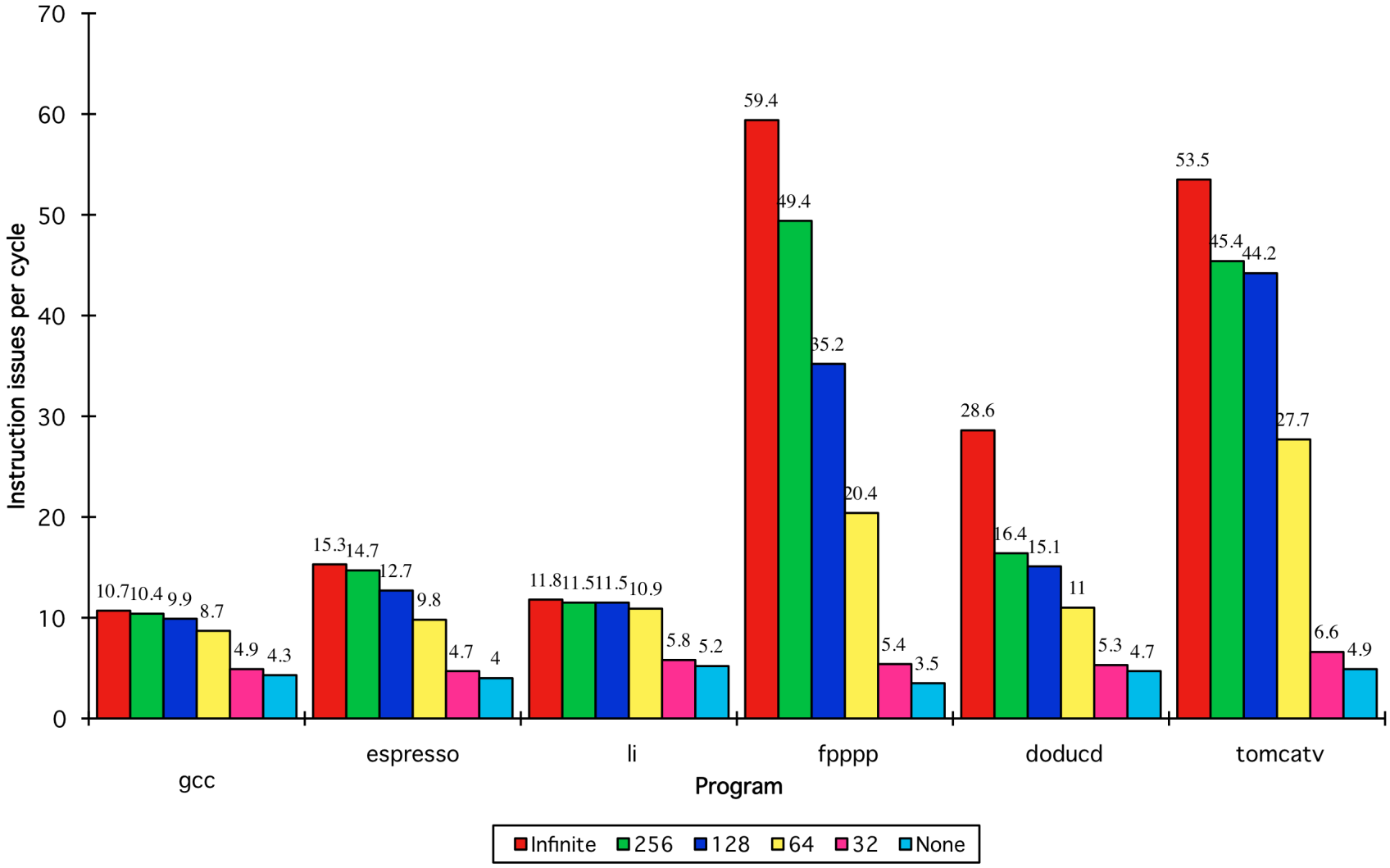
Misprediction Rates



Realistic Register Renaming

	New Model	Ideal Model	Power 5
Instructions Issued per CC	64	Infinite	4 can be issued; 6 can begin execution
Instruction Window Size	2048	Infinite	200, inc. up to 32 loads or stores
Renaming Registers	256, 128, 64, 32, none	Infinite	88 Integer + 88 FP (In addition to 64 arch regs)
Branch Prediction	8K Tournament; 2K jump, RA predictors	Perfect	2% to 6% misprediction (Tournament Branch Predictor)
Cache	Perfect	Perfect	64KI, 32KD, 1.92MB L2, 36 MB L3
Memory Alias	Perfect	Perfect	??

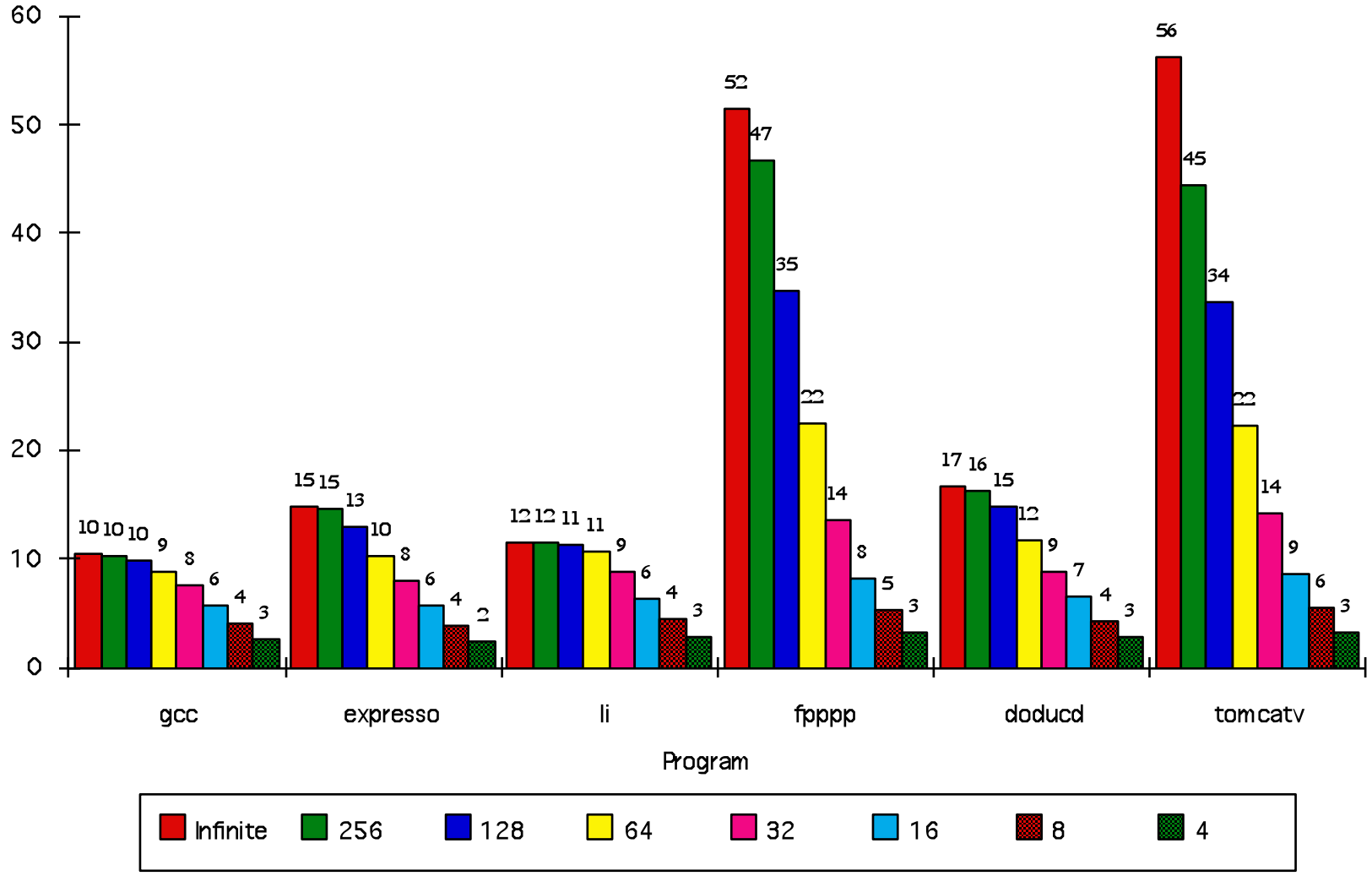
Realistic Register Renaming, Results



Realizable Hardware

	New Model	Ideal Model	Power 5
Instructions Issued per CC	64 (no restriction on instr. type)	Infinite	4 can be issued; 6 can begin execution
Instruction Window Size	Infinite, 256, 128, 64, 32	Infinite	200, inc. up to 32 loads or stores
Renaming Registers	64 Int + 64 FP	Infinite	88 Integer + 88 FP (In addition to 64 arch regs)
Branch Prediction	1K Tournament Predictor; 16-entry RA pred.	Perfect	2% to 6% misprediction (Tournament Branch Predictor)
Cache	Perfect	Perfect	64KI, 32KD, 1.92MB L2, 36 MB L3
Memory Alias	Perfect HW disambiguation	Perfect	??

Realizable Hardware, Results



Summary

- ILP is limited, even in the best case
- Two significant limiters
  - Instruction window—limits the pool of independent instructions
  - Branch prediction—limits speculation
- To continue to convert transistors into performance, new sources of parallelism!
  - Thread-level parallelism
  - Data-level parallelism

Next Time

- Multi-threading
  - Chapter 3.5