CSE240A sample questions
Administrivia

• Available on Thursday morning
  – I will upload to class website along with sending as attachment to e-mail.
• Due in a week (Dec. 8\textsuperscript{th} @ 21:59)
• Open book/open note (closed wikipedia)
• Strict page/formatting limitation

• Goes without saying: no collaboration!
Topics we’ve covered

• Technology trend / performance
• Caching / virtual memory
• ISA design
• ILP: pipelining, branch prediction
• Out-of-order execution / dataflow
• Multi-threading / multi-processing
• Heterogeneity
• Storage
Technology

• What is Moore’s law? How does it relate to processor performance? How does it not relate to processor performance?
Caching

- What is the purpose of the cache hierarchy? Why is caching superior to providing two separate memories, a small memory comprised of SRAM and a large memory comprised of DRAM, and letting the programmer choose which data to store in each memory?
Caching

• If a cache has total capacity of 24KB, is 3-way set associative, and has 16 byte lines, how many _bits_ of storage are required, in total, for the cache’s tags? Assume 64bit address space.
Caching

• Your coworker is designing a CMP and your initial idea is to have 8 cores each of which has an 8KB L1 DCache and an 8KB L1 ICache. In addition, the chip has a shared, 128KB L2. She also plan to make the L2 cache inclusive, so that any line that is in one of the L1 caches is also in the L2. You point out that this arrangement makes the L2 useless. Why? And how could you fix it?
Caching

• In your processor design, you are evaluating between having increased cache line size at L1 cache or putting a prefetcher between L1 and L2. What are the advantages of larger line size over prefetcher? What are the advantages of prefetcher over larger line size?
Caching

- TLBs are typically fully-associative while L1 caches are typically direct-mapped, or have small associativity. Why do they have different design choices in associativity when they both are among the first hardware modules to be accessed for memory?
ISA design

• Your boss just read the cryptomaniac paper and is very excited about specialized co-processors, so he gives you task of designing a specialized processor to accelerate his favorite game: Solitaire. Is this a good idea or a bad idea? How would you go about convincing your boss of this?
Pipelining / branch prediction

• Consider a simple five-stage MIPS pipeline without a branch predictor (so that the pipeline stalls until the branch resolves). You are given a really terrible branch predictor that is just 10% accurate. Would adding it to the pipeline improve or degrade the pipeline performance? Why?
Out of order execution

• Why didn’t Tomosulo suggest applying his algorithm to the integer functional units?

• How are reservation stations and register renaming related (i.e., how are their functions in the processor similar)?
Out of order execution

• Unlike IBM 360/91, HPSm has an “active instruction table” which tracks if the instructions have been completed or not andretires them in order. Why is in-order retirement of instructions important?
Dataflow

• Register renaming allows out-of-order processors to break false dependences, but the register renaming is not the only way. What is similar about how WaveScalar and TRIPS avoid false dependences? What is different?
Multi-threading

• What’s the difference between “horizontal waste” and “vertical waste” as it refers to SMT? How does SMT remove each kind of waste?
Multi-threading

• SMT (Simultaneous Multi-Threading) and Niagara’s FMT (Fine-grained Multi-Threading) both allow multiple threads to share hardware resources. For both cases, the instruction fetch stage needs to decide which thread to fetch from. How are SMT and FMT different in choosing which thread to fetch? What circumstance caused this difference?
Multi-processing

• Change in SPARC T-series design specs, why?
  – T1 = in-order 4 cores x 8 threads
  – T2 = in-order 8 cores x 8 threads
  – T3 = in-order 16 cores x 8 threads
  – T4 = out-of-order 8 cores x 8 threads