

CONFIDENTIAL

Paper 332-MEng3-Test Parallel Architectures

Department of Computing
Imperial College

Examination paper November 25, 1999

First examiner: Second examiner:

Turn over ...

- 1 In this question, consider a single-issue, dynamically-scheduled processor using Tomasulo's scheme and a Re-order Buffer (ROB) to handle speculative execution.
 - a Which parts of the processor have to monitor the common data bus (CDB)?
For the remainder of this question, suppose memory read accesses are implemented using a 'load' unit, managed in the same way as floating-point units. Thus, dynamic scheduling provides a means of accommodating unexpected memory access delays.
 - b What structural hazard(s) limit the extent to which long memory access delays can be accommodated by dynamic instruction scheduling?
 - c What data hazard limits the extent to which long memory access delays can be accommodated, even if plenty of instruction-level parallelism is present?
 - d In a machine with multiple 'load' units, memory reads may take place out of order. What problems might arise?
 - e Sketch a technique for handling speculative execution of store instructions.

(The five parts carry, respectively, 20%, 20%, 10%, 20% and 30% of the marks).

- 2a Write down (using some high-level language such as C) a simple loop operating on arrays, which suffers conflict misses with a direct-mapped cache, and with a two-way set-associative cache, but not with a three-way set-associative cache. State any assumptions you need to make, and explain your answer.
- b One proposed solution to this problem is to augment a direct-mapped cache with a “victim cache”. This is a very small, separate, fully-associative cache containing recently-displaced cache blocks. Explain how this might work.
- c Another idea is to make the two-way set-associative cache “skewed”. The proposal is to index one half of the cache using the N low-order bits of the address as usual, but to index the other half using an address which is derived by applying a pseudo-random hash function to a larger group of address bits.
Explain what the effect of this scheme would be on the loop you used in your answer to part (a) of this question.
- d A loop which runs without associativity conflicts with a direct-mapped cache is likely to suffer from some conflicts with a skewed two-way set-associative cache. Explain why this is so.

(The four parts carry, respectively, 30%, 40% and 30% of the marks).

End of Paper