UNIVERSITY OF DUBLIN

TRINITY COLLEGE

Faculty of Engineering and Systems Sciences

Department of Computer Science

B.A (Mod.) Computer Science
Junior Sophister Examination

Trinity Term 2003

3BA4 - Computer Architecture II

Thursday 22nd May SAM.BECKETT ROOM 09.30 - 12.30

Dr. J.O. Jones, Dr. Andrew Butterfield

Attempt **FIVE** questions at least two from each section.

Use separate answer books for each section.

SECTION A

Q1. What is a pipelined processor? What are the benefits of pipelining? Explain the organization & operation of the DLX five stage execution pipeline.

What are data hazards? Describe two techniques that can be used to overcome such hazards. Illustrate your answer with some simple examples.

What is a load hazard? Describe a technique that can be used to overcome such hazards. Illustrate your answer with an example.

Q2. What is a cache? How does a cache reduce the effective memory access time? Explain how a cache organisation can be characterised by the three constants LKN. Explain how a data item is searched for in an LKN cache. What special names are given to cache organisations where (i) N=1 (ii) K=1 and (iii) K=4.

Compute the number of hits & misses if the following sequence of hexadecimal addresses is applied to 64 byte 2-way set associative cache with 16 bytes per line:

0000 ⇒ 0004 ⇒ 000c ⇒ 2200 ⇒ 0014 ⇒ 000c ⇒ 2204 ⇒ 0008 ⇒
113c ⇒ 2204 ⇒ 0010 ⇒ 0030 ⇒ 1130 ⇒ 0040 ⇒ 2208 ⇒ 0008

Assume that (i) the first $\log_2(L)$ bits is used as the offset in a cache line & the next $\log_2(N)$ bits to select the set and (ii) all cache lines are initially empty/invalid. Clearly state any other assumptions made.

Q3. What is the cache coherency problem? Under what conditions are the caches in a multiprocessor system considered to be coherent?

Given a three CPU+cache multiprocessor system and the following memory requests:

CPU 0: read a2
CPU 0: write a2
CPU 0: write a2
CPU 1: read a2
CPU 1: read a0
CPU 0: write a2
CPU 0: write a2

Explain in detail what happens on each memory request (e.g. bus traffic and cache line state transitions) if (i) a write-once and (ii) a Firefly cache coherency protocol is used.

Assume (i) each cache is direct mapped with 2 cache lines (ii) even addresses map to line 0 & odd addresses to line 1 and (iii) the caches initially contain addresses a0 & a1.

Q4. What is a spin lock? Two tasks executing on separate processors in a multiprocessor update a shared variable. Why should a spin lock be used to protect access to the shared variable?

Describe the operation of the load locked, store conditional & memory barrier instructions as implemented on the Alpha microprocessor. Show how these instructions can be used to implement a simple spin lock. What are the advantage(s) of this method compared with a spin lock based on a simple test and set instruction.

SECTION B

Q5 Given the following logic function: Y = NOT(A*B+C*D)

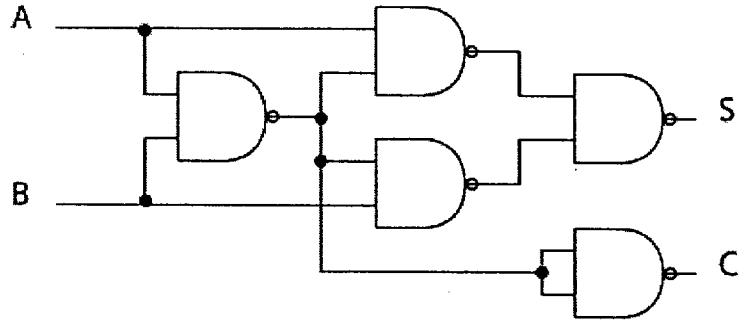
(i) Design a CMOS switch circuit that implements this function. Your design should include a determination of the widths of the transistors involved (assuming a minimum width of 0.1μm, and a μ_n/μ_p ratio of **2.5**).

(ii) Express the CMOS layout topology of your circuit using a Stick Diagram, subject to the following constraints: Power and Ground run horizontally across the top and bottom respectively, of the circuit, inputs A and C enter on the right, while inputs B and D enter on the left, on Polysilicon, and the output emerges on the top, also on Polysilicon. Your diagram should have a clear legend of the stick notation being used.

Q6 (i) A microprocessor design has a buffer that interfaces to the external databus off-chip, to handly memory accesses. This buffer consequently must drive a large capacitance load. The buffer has an output driver that is too small to drive the external signals at the desired speed. Explain why making that driver with larger transistors does not solve the problem, and show how to add in extra inverters to give a solution.

(ii) The same design requires a data-register to drive a signal to a very long internal bus line, feeding other circuits around the chip. The speed performance of this signal is important. Explain how you might re-size transistors, and add extra if necessary, in order to maximise speed of this bus line.

Q7 Context is important in IC Design - a given logic function may have different forms in various parts of an integrated circuit according to the differing relationships it has with its neighbours. Use the following design of a half-adder shown below as an example to show how some of its subparts with the *same* logic function end up having *different* physical implementations.



Q8 (i) Show how pseudo-nMOS CMOS technology can be used to produce regular programmable logic arrays (PLAs), by describing the floorplan scheme, and giving floorplan diagrams for each of the basic tiles, and a stick diagram of the array core tile

(ii) Sketch out a PLA to implement the following logic(/X denotes NOT(X)):

S = A * /B * /CI + /A * B * /CI + /A * /B * CI + A * B * CI

CO= A * B + A * CI + B * CI

where A, B and CI are inputs and S and CO are outputs. Your answer should show how the PLA building blocks are put together and the effect of the programming.