Computer Science 605.611

Problem Set 9 Answers

(Recall that by default the MIPS system uses big endian storage order for memory)

1. The addresses and 8-bit contents of four consecutive memory bytes on a system that uses big endian storage order are shown below:

Address	Contents
0x10040006	0xF1
0x10040007	0xF2
0x10040008	0xF3
0x10040009	0xF4

Register $6 contains the 32-bit address 0x10040006.

a) (3) Show (in hex 0xdddddddd) the 32-bit contents of register $2 after executing the instruction:

lh $2,2($6) $2 = _____0xFFFFF3F4___

This instruction loads a halfword from address 2+0x10040006 = 0x10040008 into the low half of register $2. The sign bit is extended throughout the upper 16 bits of $2.

b) (3) Show (in hex 0xdddddddd) the 32-bit contents of register $3 after executing the instruction:

lhu $3,0($6) $3=_____0x0000F1F2___

This instruction loads a halfword from address 0x10040006 into the low half of register $2 as an unsigned 16-bit value. The upper 16 bits of $2 are all set to 0.

2. An integer array that contains two 32-bit elements is declared as:
unsigned integer idata[2];
The element idata[0] contains the value 0x07176305
The element idata[1] contains the value 0x60151930

The memory address of element idata[0] is 0x100900C0. The address of idata[1] is 0x100900C4. Indicate in hex, the contents of the single 8-bit memory byte located at each address listed in the tables below for:

a) (8) a little endian memory system

Address	8-bit Contents
0x100900C1	0x63
0x100900C4	0x30

With little endian storage order, the bytes within a multi-byte item are ordered from low byte to high byte. The address of the low byte of idata[0] is 0x100900C0, the high byte is at address 0x100900C3. The address of the low byte of idata[1] is 0x100900C4, the high byte of idata[1] is at address 0x100900C7.
The bytes within the two elements appear in memory in the following order:

Address	8-bit Contents
0x100900C0	0x05
0x100900C1	0x63
0x100900C2	0x17
0x100900C3	0x07
0x100900C4	0x30
0x100900C5	0x19
0x100900C6	0x15
0x100900C7	0x60

b) (8) a big endian memory system

Address	8-bit Contents
0x100900C2	0x63
0x100900C6	0x19

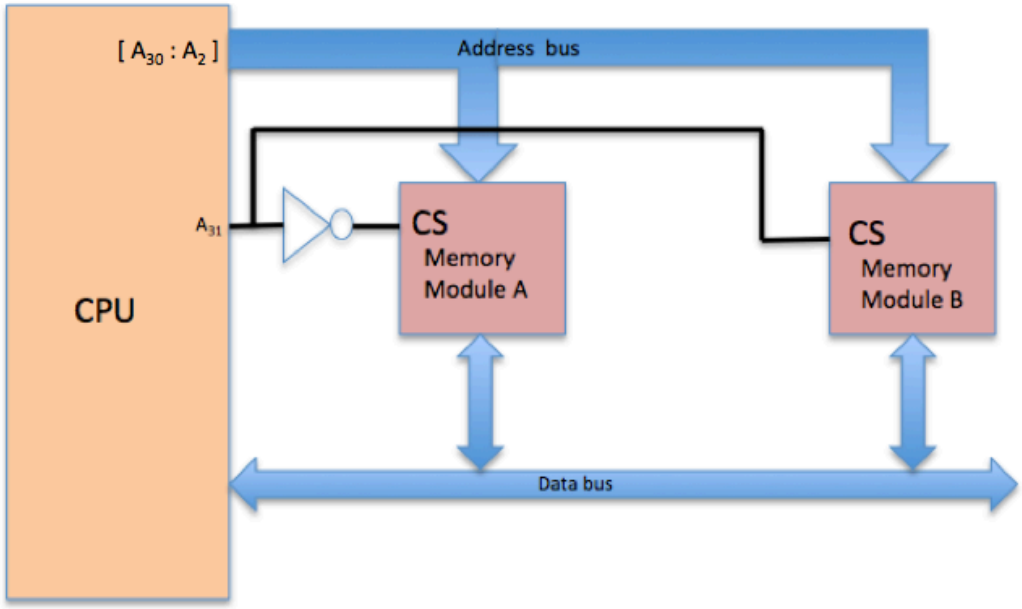
With big endian storage order, the bytes within a multi-byte item are ordered from high byte to low byte. The address of each 32-bit element is the address of the high byte within the element. The address of the high byte of idata[0] is 0x100900C0, the low byte is at address 0x100900C3. The address of the high byte of idata[1] is 0x100900C4, the low byte of idata[1] is at address 0x100900C7. The bytes within the two elements appear in memory in the following order:

Address	8-bit Contents
0x100900C0	0x07
0x100900C1	0x17
0x100900C2	0x63
0x100900C3	0x05
0x100900C4	0x60
0x100900C5	0x15
0x100900C6	0x19
0x100900C7	0x30

(The bytes within a register are always ordered from high byte on the left to low byte on the right.)

3. The memory system shown below consists of two modules (A and B). Each module has a depth of 536870912 and a width of 32. That is, each cell is 32 bits wide and there are 536870912 cells in each module. The memory storage cells within each module

are numbered starting from 0. The system employs 32-bit addresses (A31 – A0). However since each memory cell is 4 bytes wide, the 2 low bits of the memory address for each cell are always 00. So bits A1 and A0 are not used on the address bus.



The chip select (CS) is active high. That is, to select a chip, its CS input must =1.

a) (3) What is indicated by the bit pair A0 and A1 within the 32-bit memory address on this system?
These 2 bits are used as an offset to select a byte within the cell after it has been read.

b) (3) What is the minimum number of bytes transferred from memory to the CPU for a single read operation?
Each read obtains the contents of a single memory cell, which is four bytes (32 bits).

c) (3) What is the total number of bytes contained in module B?

Since the module contains 536870912 cells and each cell is 4 bytes, the total number of bytes in the module = 4 * 536870912 = 2147483648.

(3) What is the 32-bit hex address of cell 7 (the 8th cell) within module A?

A_{31} must be 0 (which is inverted and used as the chip select), bits A_{30} through A_2 must contain 7. The 32-bit address = 00000000000000000000000000011100 = 0x0000001C.

Since each cell is 4 bytes, the low two bits of the 32-bit address for each cell will always be 00.

e) (3) What is the 32-bit hex address of the next to last cell within module B?

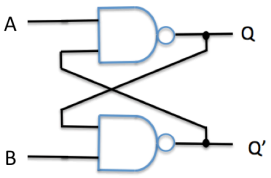
The cells are numbered 0 through 536870911 or in hex 0x0 to 0x1FFFFFFF. The next to last cell is cell 536870910 = 0x1FFFFFFE. This is the value that must be contained in the 29 bits (A_2 through A_{30}). A_{31} must be 1 to select module B. A_1 and A_0 will both be 0.
Therefore the 32-bit address of the next to last cell in module B in binary is 11111111111111111111111111111000 = 0xFFFFFFF8.

f) (4) If register $4 contains 0x80000000, which module and which byte within the module (byte 0 is the 1st, byte 1 is the 2nd byte, etc) is referenced by the instruction:

lbu $6,5($4)

The memory byte referenced is at address 0x80000005. Bit 31 = 1 so module B is selected. Bits 2 through 30 indicate cell 1 within module B. The two rightmost bits = 01 correspond to the second byte within the cell which is the 6th byte within the module.

4. (8) Complete the characteristic table for the circuit shown below by filling in the column for Q(t+1), the output in the next cycle, given that Q(t) is the output in the current cycle and the digital inputs in the current cycle are A and B. If there is no change then Q(t+1)=Q(t), otherwise Q(t+1) = 0 or Q(t+1) = 1.



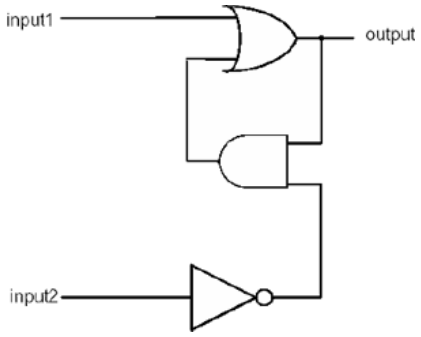
A	B	Q(t+1)
0	0	?
0	1	1
1	0	0
1	1	Q(t)

The cross coupled NAND gates within the circuit form an SR latch with active low inputs. Recall that (1 NAND X) = X', the complement of X. So if both A and B are 1, the upper NAND gate outputs the complement of Q' which is Q and the lower NAND gates outputs the complement of Q which is Q' (so the outputs retain there previous values). This means the A and B are de-asserted when each is 1 and are asserted when 0 (i.e., they are active low).

When both A and B are asserted (i.e., set to 0) each NAND gates outputs 1 (since 0 NAND X = 1). However, when A and B are de-asserted again, the final value of Q depends on whether A or B reaches 1 first. So the final output is unpredictable.
Asserting only A causes the output Q to go to 1 which then forces Q' to 0. Then de-asserting A latches or retains the current value for Q (=1).

Asserting only B causes the Q' to go to 1 which then forces Q to 0. Then de-asserting B retains the current output values (Q=0, Q'=1).

5. (5) Which one of the following (a. through e.) best describes the circuit below? Explain your answer.



a. D Latch
b. D Flip-flop
c. S-R Latch
d. S-R Flip-flop
e. Comparator

A clock does not control when the output changes so this is not a flip flop.

In the zero state output=0 by definition. In the 1 state output=1. For state 0, if both input1 and input2 are 0, the AND gate generates (output AND input2') = 0 AND 1 = 0, so the OR gate outputs 0 OR 0 = 0. Thus the zero state is retained.

In the zero state, if input1 = 1 and input2 = 0, the OR gate outputs 1 (the state changes to 1). The output remains 1 even when input1 is then de-asserted since the AND gate still produces (1 AND 1) = 1 [i.e.. output AND (input2)']. Thus the one state is retained.

In the one state, asserting input1 leaves the output at 1 but asserting input2 forces the output to 0. This is because the inverter outputs (input2)' = 0. This forces the AND gate to output 0 which causes the OR gate to generate output = 0 OR 0 = 0. So input1 acts like a SET signal and input2 acts like a RESET signal and the circuit behaves like an S-R latch.

If both input1 and input2 are asserted (i.e., set to 1), the AND gate outputs 0 and the OR gate outputs 1. When they are de-asserted to retain the output, the output stays at 1 if input2 reaches 0 first; but the OR gate outputs 0 if input1 reaches 0 first. So the output is unpredictable just like for an S-R latch.

The circuit is not a comparator since the output can be 0 or 1 when the inputs are equal or when they differ.

6. Recall that performing a read from a dynamic RAM (DRAM) requires that the chip pre-charge before it can supply the requested data. Assume that it takes 10ns to pre-charge and 15ns to either output the requested data in response to a read operation or to store the input data for a write operation. Also recall that our MIPS pipeline system employs a Harvard Architecture, transfers 32 bits at a time between the CPU and memory, and each pipeline stage consumes one clock cycle.

   a) (5) If the memory stage is the most time consuming stage, what is the maximum clock rate (expressed in GHz) that can be used for our 5-stage pipelined system using this type of DRAM memory?

The maximum clock rate would be determined by the longest stage (the memory stage). The cycle time must be at least as long as the memory cycle time = 10ns + 15ns = 25ns. This corresponds to a maximum clock rate of 1/25ns = $0.04 * 10^9$ Hz = 0.04 GHz.

   b) (5) The instruction memory contains one machine instruction per cell and is implemented as a single memory module with a storage capacity of 268435456 bytes, what is the width and depth of the module?

The width of each storage cell is 32 bits (4 bytes) which is the size of a machine instruction. The depth of the memory module is the number of cells in the module = $(256* 2^{20}) /4 = 64 * 2^{20} =$ 67108864.

7. The diagram below shows the 32-bit address format used for a particular byte addressable memory system containing 16 modules:

31 28	27 1	0
Module#	Cell number	offset within cell

a) (3) What is the width in bits of each memory module? Width = _16_ bits.
A single-bit offset corresponds to 2 bytes per cell (i.e., 2*8 = 16 bits)

b) (4) How many memory modules must be accessed to read a single 32-bit word from address 0x00408000 with this memory system?
Since the module number appears in the leftmost field, this is a high-order interleaved system with adjacent cells within the same module. So a single module has to be accessed twice to obtain a single 32-bit word from address 0x00408000.

c) (4) How many memory modules must be accessed to read a single 32-bit word from address 0x3FFFFFFE with this memory system?
This address corresponds to the final cell within module 3. So two modules must be accessed to obtain a single 32-bit word from this address (the final cell in module 3 and the first cell in module 4).

8. (9) Our byte addressable MIPS system employs a 32-bit CPU-to-memory bus as well as memory cells that are 32 bits wide. Recall that an aligned memory access is one for which the

memory address used for the read or write is a multiple of the size (in bytes) of the item that is being transferred.
Ignoring any exceptions that may occur, how many data memory cells must be read by each of the following instructions if the contents of register $8 = 0xB003D8E0 ?

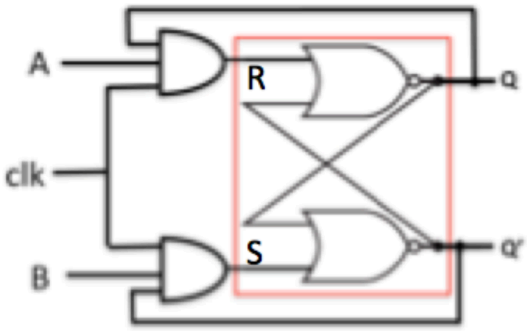
I. lb $t0,54($8)
Only one memory cell is read because the byte resides within a single 32-bit cell.

II. lh $t0,54($8)
Decimal 54 = 0x36 so the address is 0xB003D8E0+0x36 = 0xB003D916 which falls within the word at address 0xB003D914 (containing the 4 bytes at addresses 0xB003D914, 0xB003D915, 0xB003D916 & 0xB003D917 respectively). The two bytes at addresses 0xB003D916 and 0xB003D917 reside within the same cell, so only one cell is read.

III. lw $t0,54($8)

The address is 0xB003D8E0+0x36 = 0xB003D916. This address falls within the word whose bytes are at addresses 0xB003D914, 0xB003D915, 0xB003D916, and 0xB003D917. Hence the first 2 bytes are in this word and the next 2 bytes are in the next adjacent word whose bytes are at addresses 0xB003D918, 0xB003D919, 0xB003D91A and 0xB003D91B. Hence two cells must be read to obtain the required 4 bytes (the final 2 bytes of one cell and the first 2 bytes of the next cell). An unaligned exception occurs to signal the problem.

9. (8) Complete the characteristic table for the circuit shown below by filling in the column for Q(t+1), the output in the next cycle, given that Q(t) is the output in the current cycle and the digital inputs in the current cycle are A and B which take effect when the clock signal (clk) pulses from low to high and back to low. The width of the pulse is such that the outputs Q and Q' only change once. If there is no change then Q(t+1)=Q(t), otherwise Q(t+1) = 0 or Q(t+1) = 1.



A	B	Q(t+1)
0	0	Q(t)
0	1	1
1	0	0
1	1	Q(t)'

This circuit is just the NOR gate implementation of an S-R latch. But here the upper AND gate generates the RESET signal (R) and the lower AND gate generates the SET signal (S). Both AND gates output 0 when the clk is low so there is no change in the outputs Q and Q'.

When the clk signal goes from low to high, the output of the upper AND gate depends on A and Q; the output of the lower AND gate depends on B and Q'.

When A is 1 and Q is 1, the output of the upper AND gate (R) is 1 (a reset operation)
When B is 1 and Q' is 1, the output of the lower AND gate (S) is 1 (a set operation)

If A and B are both 1 at the same time, the upper AND gate outputs 1 AND Q = Q, and the lower AND gate outputs 1 AND Q' = Q'. So if Q is 1, R= 1 AND 1 = 1 (Q is reset). If Q=0 (i.e., Q' is 1), S= (1 AND 1) = 1 (Q is set). Hence Q(t+1) = Q(t)' when both A and B are 1 (the output toggles).

A circuit called a J-K flip flop behaves in this way. That is, at the clock edge J-K flip flop toggles the output if the inputs J and K are both 1, the output is unchanged if J and K are both 0, the output is reset to 0 if J=1 and K=0, the output is set to 1 if J=0 and K=1.

10. A computer system designed and built in the 1970's made use of a single 8-bit CPU-to-memory data bus and a single separate address bus. Suppose that such a system contained 4 memory modules or chips, each of which was 8 bits wide with a depth of 16777216. A memory access for each module consisted of only two phases: an addressing phase in which an address is sent to select a specific memory module and cell within the module; and a data transfer phase in which the selected data item is read from or written into. Each phase takes 40 nano-seconds to complete. What is the minimum time (in nano-seconds) required to read a single 32-bit data item from this memory system if it employs:

a) (4) high order interleaving

Since the data bus is only 8 bits wide, only a single byte can be transferred over the bus at one time.

With high order interleaving, consecutive or adjacent cells reside within the same module. So a separate addressing phase and data transfer phase is required for each byte. The minimum time required to read a single 32-bit data item is therefore 4*(40+40) = 320 ns.

b) (4) low order interleaving

With low order interleaving, consecutive cells reside within different modules.

Hence the modules can operate in parallel so that while a data transfer occurs for one module the addressing phase for another module can occur in parallel.

Therefore the first byte is obtained after 40+40 = 80 ns and the addressing phase for the second module overlaps with the data transfer phase of the first, etc. So the entire group of 4 cells required for the complete 32-bit data item is obtained in

80 + 40 + 40 + 40 = 200 ns.