

Name: Sean Penney

CS 472 Midterm 2

ID#: 931-650-344

*All answers should be typed directly below the questions. All work should be shown. Answers with no work shown will be considered cheating. The exam should be typeset via L<sup>A</sup>T<sub>E</sub>X. All questions should be included in the exam.*

*This is an open book exam. As it is a take home exam, you are on the honor code. Use your book, use your notes, don't use your friends.*

*There are 200 points possible on this exam.*

*Be detailed in your answers. Remember, this is a take home exam, so implied answers will not be considered. Take the time and make everything explicit!*

1. (Bonus)

Turtle.

Yeah Yeah Yeah Yeah

Miami, uh, uh

Southbeach, bringin the heat, uh

Haha, can y'all feel that

2. (10 points) 6.7

a.  $50 + 50 + 50 = 150$  ns.

b.  $40 + 40 + 40 = 120$  ns.

c. Cycle 3 has to be replaced by two clock cycles.

$30 + 30 + 30 + 30 = 120$  ns.

d. Cycles 2 and 3 have to be replaced by two clock cycles each.

$20 + 20 + 20 + 20 + 20 = 100$  ns.

e. Cycle 2 has to be replaced by 3 clock cycles. Cycle 3 has to be replaced by 4 clock cycles.

$10 + 3*10 + 4*10 = 80$  ns.

3. (10 points) 6.13

Original CPI =  $.65*1 + .1*5 + .05*2 + .2*8 = 2.85$

Want CPI to be 2.28 (20% reduction). This is .57 less than the original

This could be achieved by having 5 cycles per conditional branch instruction.

4. (10 points) 6.30

The arithmetic mean is simple to calculate, but one outlier can throw off the result.

One advantage of the geometric mean is that it doesn't matter if the geometric mean of the unnormalized data is taken then normalized, or vice versa. A disadvantage is that the geometric mean is not proportionate.

The harmonic mean can be used to average a series of tests that yield rates.

5. (10 points) 6.31

$$(\sqrt{x} - \sqrt{y})^2 \geq 0$$

$$x - 2\sqrt{xy} + y \geq 0$$

$$x + y \geq 2\sqrt{xy}$$

$$\frac{x+y}{2} \geq \sqrt{xy}$$

6. (10 points) 6.32

If absolute times are used, rather than normalized times, the execution time may be dominated by a test program with a large data set.

Disadvantages are that an old system might be used as the reference, or a system that performs well with a particular task.

Another problem is that units are lost when values are normalized with respect to the reference machine.

7. (10 points) 7.13

Between the instruction memory block and the register file block, there would need to be logic to figure out whether the branch can be predicted with confidence. If the branch is hard to predict, the hardware should execute the predicated code by calculating the appropriate address and sending it to the program counter. If the branch can be predicted, then the expected direction of the branch should be taken and the appropriate address should be calculated.

8. (10 points) 7.19

a. Larger numbers of registers is not always better. With more registers, more work needs to be done to store and restore when calling and returning from functions. Also, a larger register file requires more complexity and resources. Assuming the processor has an adequate cache, data can be accessed quickly from there.

b. It is beneficial for new CPUs to be software-compatible with old CPUs, which requires the new CPU to have no additional registers the programmer can use. Doubling the number of general-purpose registers requires adding another bit to each instruction.

c. General purpose registers give the programmer flexibility. General purpose registers are less efficient though, since the registers need to be marked to signify what they are for.

d. More registers could become dedicated registers, so that the number of general purpose registers remains the same.

9. (10 points) 7.20

UNIX and RISC both provide simple tools that make the system better. The idea behind RISC is that simple instructions leads to higher performance, by finding ways to run the individual instructions much faster. The UNIX philosophy is focused on providing simple tools, a unified filesystem and shell scripting to combine these tools.

10. (10 points) 9.12

a. Using a direct mapped cache, the data is stored like a simple table, with columns for data and tags.

b. A fully associative cache allows you to store the data anywhere in the cache.

c. A set-associative cache is a hybrid between a fully associative and direct mapped cache. There is a general location in which the data is to be stored, and within that general location the data can be stored anywhere.

11. (15 points) 9.21

A physical data cache has longer access time since the MMU first needs to perform a logical-to-physical address translation. A logical cache does not need to wait for an address translation. The disadvantage of a logical cache is that when the logical-to-physical address mapping is modified, the data in the cache cannot be used and the logical cache has to be flushed.

12. (10 points) 9.23

The average access time of a system with a cache that's accessed in parallel with main store is  $t_{ave} = ht_c + (1 - h)t_m$

Let  $t_1$  be the time to fetch a line from main store to reload the cache on a miss. We must add  $(1 - h)t_1$  to the average access time.

$$t_{ave} = ht_c + (1 - h)t_m + (1 - h)t_1$$

13. (10 points) 9.26

$$\begin{aligned} t_{ave} &= h_1t_{c1} + (1 - h_1)h_2t_{c2} + (1 - h_1)(1 - h_2)t_m \\ &= 2.22 \text{ cycles} \end{aligned}$$

14. (10 points) 9.35

There are 4 sets, so 2 bits are required for the set since  $2^2 = 4$

Each set is 2 MB in size, and there are 32 bytes per line.

2 MB / 32 bytes =  $2^{16}$  bytes. Thus, 16 bits are needed for the line.

Each line has 4 words, so 2 bits are needed for the offset, since  $2^2 = 4$ .

15. (10 points) Assume a 64-bit virtual address and a 64-bit physical address. The page size is 4KB. How many total entries are there in the page table? Express your answer in powers of 2.

$$\text{Number of entries} = \text{size of memory} / \text{page size} = 2^{64} / 2^{12} = 2^{52}$$

16. (15 points) Describe the ARM pipeline. What benefit does this serve over a single cycle flow through processor?

The ARM7 has a three-stage pipeline, which increases instruction flow up to three times. Instead of each instruction executing one at a time, the steps of several instructions are executed at the same time. The stages are fetch, decode, and execute, and branches are improved by calculating the PC 8 bytes ahead of the current instruction. Other ARM designs have more stages, and better branch prediction. This allows for instructions to be executed faster than a single cycle flow through processor. Branch prediction helps so that cycles are not lost by flushing the pipeline.

17. (15 points) Describe a method of programmatically determining cache line size as well as cache sizes.

A large char array that exceeds the cache size should be allocated and filled with data. Values should be retrieved in step sizes of  $n$  bytes, and some operation should be performed on them. Have a timer that keeps track of how long it takes to retrieve the  $n$  bytes and perform an operation. When  $n$  reaches the cache line size, the time to retrieve and perform operations on  $n$  bytes should sharply increase.

A similar approach can be taken to determine the cache sizes. Allocate a large array, and time access to chunks of the array. When the time to access a chunk of the array increases, the size of the chunk is the size of the cache.

18. (25 points) On modern CPUs, there are often multiple cores. Describe ways in which you believe cache design had to change to deal with this situation.

Issues related to cache coherency need to be addressed when using multiple cores. Say two cores have the same data in their L1 cache. At first, the data is the same for both cores. If processor 0 performs a store operation that modifies the data, then processor 1 performs a load, the load must somehow see the new value. As soon as processor 0 requests to write to the cache, the copy of data in processor 1's cache must be removed. The processor 1 would then have a cache miss, and reload the new data.