EECS 370
Final Exam

Name:

Open book, open notes. Calculators are permitted, but no laptops, PDAs, cell phones, etc. This exam has 12 pages, 9 questions, and 100 points. Questions vary in difficulty; it is strongly recommended that you do not spend too much time on any one question.

1. *(10 points)*: Please circle TRUE or FALSE for each of the following questions.

(a) FALSE
Doubling the block size of a cache cannot affect its compulsory miss rate.

(b) FALSE
A multi-level page table saves memory because the page table is shared by all process contexts.

(c) TRUE
A virtually indexed caching scheme, while fast from a circuit perspective, complicates context switches and memory sharing.

(d) FALSE
Adding simultaneous multithreading (SMT) to a processor cannot improve pipeline CPI unless applications are recoded to take advantage of SMT features.

(e) TRUE
The Z-buffer in a graphics processing unit allows polygons to be drawn on the screen in any order.

2. *(13 points)*: For all questions, circle the BEST choice from the list provided.

Memory system performance can sometimes be improved by having the cache set index consist only of address bits that are not affected by virtual to physical address translation.

2.1: Why is this?

(a) This configuration will reduce the cache miss rate.

(b) This configuration will reduce the page fault rate.

(c) This configuration will reduce the TLB miss rate.

(d) This configuration will reduce the amount of memory needed for page tables.

(e) This configuration will require fewer comparators to check for tag matches.

*** (f) This configuration will allow the data and tags from the selected cache set to be retrieved at the same the time that address translation is being performed.

You decide that the next computer you build is going to have this property as described above. Some parameters of your system design are as follows: 32 bit addresses, byte addressable memory, 4 kilobyte page size, 2-way set associative cache, 32 byte cache block size, write-allocate write-back policy.

2.2: Given the parameters above, what cache sizes will have this property (that the set index bits consist entirely of bits that are not affected by the virtual to physical address translation)?

(a) Only caches with exactly 4096 bytes of data

(b) Any cache with 4096 bytes or more of data

(c) Any cache with 4096 bytes or less of data

(d) Only caches with exactly 8192 bytes of data

(e) Any cache with 8192 bytes or more of data

*** (f) Any cache with 8192 bytes or less of data

(g) All caches will have this property

(h) No caches will have this property

Assuming all other variables except for the specific ones below are unchanged, and you have the single goal of building a cache with the largest possible size (in total data bytes) that has the property described on the previous page, select the best choice for each parameter described.

2.3: Cache block size

    (a) Use a larger cache block size

    (b) use a smaller cache block size

    *** (c) Cache block size make no difference toward your goal

2.4: Degree of associativity for the cache

    (a) Use a lower degree of cache associativity

    *** (b) Use a higher degree of cache associativity

    (c) Cache associativity makes no difference toward your goal

2.5: TLB size:

    (a) Use a larger TLB

    (b) Use a smaller TLB

    *** (c) TLB size makes no difference toward your goal

2.6: Degree of associativity for the TLB:

    (a) Use a lower degree of TLB associativity

    (b) Use a higher degree of TLB associativity

    *** (c) TLB associativity makes no difference toward your goal

2.7: Page size:

    *** (a) Use a larger page size

    (b) Use a smaller page size

    (c) Page size makes no difference toward your goal

2.8: Number of page table levels:

    (a) Use a single level page table

    (b) Use a two level page table

    (c) Use a three level page table

    *** (d) Number of levels of the page table makes no difference toward your goal

3. *(8 points)*:

Consider a cache as follows:

- Byte addressable

- Direct mapped

- 8 bytes total cache data size

- 2 byte cache block size

A sequence of eight memory reads is performed in the order shown from the following addresses:

0, 11, 4, 14, 9, 1, 8, 0

(a) For the given cache, how many of these references result in cache misses?

ANSWER: 8

(b) How many of them are compulsory misses?

ANSWER: 5

(c) How many of them are capacity misses?

ANSWER: 1

(d) How many of them are conflict misses?

ANSWER: 2

4. *(12 points)*:

A standard IEEE single precision floating point multiplication between two registers is performed.

(a) The first floating point register contains the hexadecimal bit pattern: 0xc0d00000 (binary 11000000110100000000000000000000). This bit pattern is representing a floating point number in the standard IEEE format. State the contents of this register as either a decimal number or a fraction.

ANSWER: -6.5 or $-\frac{13}{2}$

(b) The second floating point register contains the fraction 5/8 (0.625 decimal). State the 32 bit floating point value in this register in either binary or hexadecimal.

ANSWER: 3f200000 or 00111111001000000000000000000000

(c) After the multiplication, what is the representation of the product? State the 32 bit floating point product in either binary or hexadecimal. For full credit, you must show your work for multiplying the two numbers in IEEE floating point format. Converting to decimal and multiplying in decimal will receive only partial credit.

ANSWER: c0820000 or 11000000100000100000000000000000

5. *(14 points)*:

Consider a superscalar LC2 constructed as follows:

There are two pipelines, and each pipeline has five stages (similar to the lecture slide pipelined LC2). The pipelines are named Pipe-0 and Pipe-1, and their respective stages are named: IF0, ID0, EX0, MEM0, WB0, IF1, ID1, EX1, MEM1, WB1. Similarly, the pipeline registers are called: IFID0, IDEX0, EXMEM0, MEMWB0, IFID1, IDEX1, EXMEM1, MEMWB1. Both pipelines have full access to all resources, including memory; there are no structural hazards. Like the lecture slide pipeline, registers are written in WB in the first half of the clock cycle and read in ID in the second half of the clock cycle; thus, no WBEND forwarding is required. Instructions are always fetched in aligned pairs. In other words, a pair of instructions are fetched from addresses X and X+1, where X is even and X+1 is odd. The instruction from X ends up in IFID0 and the instruction from X+1 ends up in IFID1. Full forwarding is available between all stages of both pipes. If a data hazard is detected that cannot be resolved by forwarding, both pipes stall their instruction fetches and one NOOP is inserted into an appropriate place in each pipe. If one instruction of a pair is not to be executed as a consequence of a branch, it is squashed by replacing it with a NOOP.

(a) For each **add** instruction in the following assembly language code, fill in the two blanks next to it with the name of pipeline register that supplies each of the two operands to the ALU for that add. This code begins at address zero. Possible choices for your answers are: IFID0, IDEX0, EXMEM0, MEMWB0, IFID1, IDEX1, EXMEM1, MEMWB1.

add	2 3 1	r2 from:	IDEX0	r3 from:	IDEX0
add	4 5 2	r4 from:	IDEX1	r5 from:	IDEX1
add	1 2 3	r1 from:	EXMEM0	r2 from:	EXMEM1
add	2 1 4	r2 from:	EXMEM1	r1 from:	EXMEM0
add	1 3 5	r1 from:	MEMWB0	r3 from:	EXMEM0
add	2 4 6	r2 from:	MEMWB1	r4 from:	EXMEM1
halt					
noop					

(b) Do any stalls occur in the above sequence? If so, indicate the pair(s) of instructions with dependencies between them that require stalls.
ANSWER: NO

6. *(12 points)*: Consider the following computer with a CPU, data cache, and memory.

- The system has a 12-bit virtual address size, and the size of a virtual memory page is 256 bytes. The computer has 65536 bytes of physical memory.

- The size of the CPU data cache is 512 bytes, with a block size of 16 bytes. The CPU data cache is 2-way set-associative with LRU replacement, and it is a virtually indexed cache. For a hit, the data cache can be accessed in a single clock cycle. For a miss, a cache block in memory can be read in 10 clock cycles, which includes the time for a TLB access.

- The CPU includes a fully associative TLB with 2 entries and an LRU replacement policy. For a TLB miss, the appropriate page table entry is supplied to the CPU after 25 clock cycles.

The system uses a single level page table. Assume the TLB and the data cache are initially empty. The contents of the page table are shown below:

VPN	Valid	PPN
0	0	-
1	1	0xfd
2	1	0x4b
3	0	-
4	1	0x55
5	1	0x32
6	1	0x54
7	0	-
8	0	-
9	1	0xfe
a	1	0xfc
b	0	-
c	1	0xba
d	1	0x01
e	0	-
f	0	-

(continued from the previous page)

(a) Data reads from the following virtual addresses are performed (in the order listed):
0xce2, 0x258, 0xcea, 0x2d4, 0x15b, 0x241, 0xcee.
Please complete the table below to show the following for each read:

- Physical address (show it in hex)

- Data cache hit/miss (write **HIT** or **MISS** as appropriate)

- TLB hit/miss (write **HIT** or **MISS** as appropriate, or write **NA** if the TLB is not accessed)

Please show your work to be eligible for partial credit.

Virtual address read	Physical address (in hex)	Data cache hit/miss	TLB hit/miss/NA
0xce2	bae2	MISS	MISS
0x258	4b58	MISS	MISS
0xcea	baea	HIT	NA
0x2d4	4bd4	MISS	HIT
0x15b	fd5b	MISS	MISS
0x241	4b41	MISS	HIT
0xcee	baee	HIT	NA

(b) A program running on the above described computer has a data cache miss rate of 10% and a data TLB miss rate of 5%. What is the average memory access latency for this program for data accesses? Please show your work and clearly state any assumptions you make.

ANSWER: 1 + 0.1 * 0.95 * 10 + 0.1 * 0.05 * (25 + 10) = 2.125 cycles

7. *(9 points)*: A certain not-so-new, high-performance computer has a cache with the following characteristics:

- 36-bit addresses

- Byte addressable memory

- 4MiB cache size (4MiB = 2^{22} bytes)

- 8KiB cache block size (8KiB = 2^{13} bytes)

- 32-way set associative cache

On this computer, addresses bits are numbered as follows:

- Least significant bit: Bit 0

- Most significant bit: Bit 35

Fill in the following table to show how address bits are used with this cache.

- For the Size column, indicate how many bits this field contains.

- For the Range column, indicate the range of bits that make up this field. For example, the least significant five bits would be bits 0-4, and the most significant three bits would be bits 33-35.

- For the Value column, indicate the specific value for this field in either hexadecimal or binary for the address 0xadf85458a

Field	Size (Number of bits)	Range (Bit x - bit y)	Value (hex or bin) for address 0xadf85458a
Block Offset	13	0 - 12	58a or 0010110001010
Set Index	4	13 - 16	a or 1010
Tag	19	17 - 35	56fc2 or 1010110111111000010

8. *(12 points)*: A certain computer has a virtual memory system configured as follows:

- 32-bit addresses, byte addressable
- 2KiB page size (2KiB $= 2^{11}$ bytes)
- All page table entries are 4 bytes
- Can use either a 1-level or 3-level page table.
- If a 3-level page table is used, 7 address bits are devoted to each level.

Please answer the following questions:

(a) For the 1-level page table, how many page tables are there if all addresses are used?

ANSWER: 1

(b) For the 1-level page table, how much memory is used for page table(s) if all addresses are used?

ANSWER: 2^{23} bytes or 8 MiB

(c) For the 1-level page table, how many page tables are there if only the first 4,096 and last 4,096 addresses are used?

ANSWER: 1

(d) For the 1-level page table, how much memory is used for page table(s) if only the first 4,096 and last 4,096 addresses are used?

ANSWER: 2^{23} bytes or 8 MiB

(e) For the 3-level page table, how many page tables are there if all addresses are used?

ANSWER: $1 + 2^7 + 2^{14} = 16{,}513$ page tables

(f) For the 3-level page table, how much memory is used for page table(s) if all addresses are used?

ANSWER: $16513 * 2^9 = 8{,}454{,}656$ bytes or 8256.5 KiB

(g) For the 3-level page table, how many page tables are there if only the first 4,096 and last 4,096 addresses are used?

ANSWER: 5

(h) For the 3-level page table, how much memory is used for page table(s) if only the first 4,096 and last 4,096 addresses are used?

ANSWER: $2^7 * 4 * 5 = 2{,}560$ bytes or 2.5 KiB

9. *(10 points)*:
A certain Fujitsu disk has the following characteristics:

- 384,000 tracks (total for all cylinders)

- average 748 sectors/track

- 4 double-sided platters, 8 heads

- standard sector size (512 bytes)

- Average wait time to access disk: 20 milliseconds

- Seek time (average) 4.5 milliseconds

- Seek time (track to an immediately adjacent track) 0.3 milliseconds

- Rotational Speed: 10,000RPM

- Disk can read from all heads at once

Show your work for partial credit.

(a) What is the capacity of this disk? Round your answer to the nearest GB. (1GB $= 10^9 =$ 1,000,000,000 bytes).

ANSWER: 384000 * 748 * 512 = 147,062,784,000 bytes.
Rounding to the nearest GB: 147 GB

(b) Assume files are laid out optimally with no fragmentation. Each file uses all tracks on a single cylinder. If a file is too large to fit on a single cylinder, it continues onto an immediately adjacent cylinder to minimize seeking. What is the average time required to service a request to read a 10MiB file from disk? (1MiB $= 2^{20}$ bytes)

ANSWER: $20 + 4.5 + (0.5 + \frac{10*2^{20}}{748*8*512}) * \frac{1000*60}{10000} + 3 * 0.3 =$ 48.9 milliseconds