Seat No. _____

King Mongkut's University of Technology Thonburi

Final Exam of First Semester, Academic Year 2015

CPE 325 Computer Architectures and Systems

Computer Engineering Department, 3rd Yr.
Section: ABCD

Wednesday 25th November 2015

9.00-12.00

Instructions

1. This 10-page exam contains 3 parts (25, 3, and 6 questions, respectively) for a total of 100 points.
2. The answers must be written in this exam paper. Please read the instructions carefully.
3. A calculator and a paper dictionary are allowed.
4. A single A4-sized note may be taken into the examination room. The note has to be handed in with the exam.

Students will be punished if they violate any examination rules. The highest punishment is dismissal.

This examination is designed by
Assoc. Prof. Tiranee Achalakul, Ph.D.
Asst. Prof. Marong Phadoongsidhi, Ph.D.
Tel. 081-922-8466

Name:_____ Student ID:_____ Section:____

Instruction: This exam has 3 parts for a total of 100 points. You have 3 hours to work on the problems. Please write your NAME, ID, Section on EVERY page of the exam. You are allowed to bring with you (1) one A4 sheet of notes, (2) a calculator, and (3) a paper dictionary. Hand in your note sheet with the exam before leaving the room.

Part I (50 points): students will get 2 points for each right answer and *-1 point* for each wrong answer. Each question only has one right answer. Use the answer sheet in the last page of this part for your answers.

1. The main purpose of having memory hierarchy is to
    a. Reduce an access time
    b. Provide large capacity
    c. Reduce propagation time
    d. Both a and b
2. Which statement is true?
    a. DRAM is a new and fast RAM technology
    b. SRAM is smaller than DRAM
    c. RAM is faster than ROM
    d. Cache is usually built using the SRAM technology
    e. b and d
    f. a and c
    g. all of the above
3. The reason (s) for the implementation of the cache memory is
    a. To increase the internal memory of the system
    b. To offset the difference in speeds of operation between the processor and memory
    c. To reduce the averaged access and cycle time
    d. To decrease the overall cost of the computer
    e. a and b
    f. b and c
    g. all of the above
4. The temporal aspect of the locality of reference means
    a. That the recently executed instruction will be executed again soon
    b. That the recently executed instruction is temporarily not referenced
    c. That the instructions in the same sub-routine will likely be executed close to one another
    d. That the recently executed instructions may not be executed again soon
5. The correspondence between the main memory blocks and those in the cache is given by
    a. Mapping function
    b. Associativity function
    c. Replacement function
    d. Hash function

Name:_____ Student ID:_____ Section:____

6. The write-through procedure is used
    a. To write onto memory directly
    b. To write and read from memory simultaneously
    c. To write directly onto both memory and cache
    d. To write in cache
    e. None of the above
7. The bit used to signify that the cache location is updated is
    a. Dirty bit
    b. Update bit
    c. Valid bit
    d. Reference bit
    e. Flag bit
8. The drawback of using the early start protocol is
    a. Time delay
    b. Complexity of the circuit
    c. Latency
    d. High miss rate
    e. a and b
9. For direct mapping, in a 16-bit address system, if the cache is 1K in size, the tag field has _____ bits.
    a. 12
    b. 8
    c. 10
    d. 4
    e. 6
9. The technique of searching for a block by going through all the tags is called _____
    a. Linear search
    b. Binary search
    c. Associative search
    d. Non-linear search
10. For 4 ways set-associative mapping, in a 16-bit address system, if the cache is 1K in size, the tag field has _____ bits.
    a. 12
    b. 8
    c. 10
    d. 4
    e. 6
11. The extra time needed to bring the data in from the main memory in case of a cache miss is called _____
    a. Propagation time
    b. Delay time
    c. Miss Penalty
    d. Miss time
    e. Miss rate

Name:_____ Student ID:_____ Section:____

12. What is this main purpose of memory interleaving?
    a. To provide an effective way to parallelize memory access
    b. To reduce the cache access time
    c. To reduce the cache miss penalty
    d. To reduce memory access latency
    e. a and c
    f. b and c
    g. all of the above
13. Identify the correct statement about virtual memory organization
    a. It is controlled by the memory management unit
    b. It allows a large program or a program with large data to run more efficiently
    c. It can provide a faster memory transfer
    d. It is an effective way to allow multiple programs to be executed at the same time
    e. a and d
    f. b and c
    g. b and d
    h. all of the above
14. Assume that a computer system has a 32-bit address, a 4KB cache, and 1work / block. What are the number of bits needed for tag, block, and offset?, if "direct mapping" is used.
    a. 20, 10, 2
    b. 22, 10, 0
    c. 20, 12, 0
    d. 18, 12, 2
15. Which statement is false?
    a. A larger block size will increase the miss penalty
    b. CPU pipeline will be stalled during a cache miss
    c. LRU (Least recently used) is a typical replacement algorithm used in a direct map cache
    d. a and c
    e. a and b

Use the following paragraph to answer questions ***16-18***
A data block can be read from a main memory to cache with the following specification: 1 cycle to transfer an address; 20 cycles to access each word; 1 cycle to transfer each word to cache. In addition, the system also has two separate caches (Instruction and Data), where the I-cache and D-cache miss rate = 2%; Base CPI = 2; and Load & stores are 30% of instructions.

16. In a system with 8-word block, 1-word-wide DRAM with 4-way memory interleaving, how many cycles are needed for a miss penalty?
    a. 28
    b. 29
    c. 36
    d. 49
    e. 56

17. If the above system has no memory interleaving, the miss penalty will be _____
    a. About the same
    b. Approximately 2 times larger
    c. Approximately 3.5 times larger
    d. Approximately 4 times larger
    e. Approximately 5.5 times larger
18. With no memory interleaving, what would be the average memory access time (AMAT) of this system?
    a. 6.4
    b. 5.8
    c. 5.44
    d. 4.6
    e. 2
19. Which statement is/are true?
    a. The miss penalty directly affects the average number of clock cycles needed per instruction.
    b. Memory interleaving improves the cache access time
    c. Block transfer time can be reduced by memory interleaving
    d. All of the above
    e. None of the above
20. Which statement is false?
    a. Multi-level caches can be used to optimize the overall memory performance
    b. In comparison to L1 cache, L2 cache focuses more on low miss rate
    c. L1 cache is always faster than L2 cache
    d. L2 cache is smaller than L1 cache
21. From the following statements
    1. Lookup a page table in order to get the physical page number
    2. Take the content of the page table register and use it to identify the location of a page table
    3. Translate a virtual page number into a physical page number
    4. Copy the page offset into physical address
    5. Send the physical address to the cache controller to obtain the data
    6. TLB miss occur
    7. CPU send the virtual address to TLB

    What are the correct sequence of actions?
    a. 7, 6, 2, 1, 3, 4, 5
    b. 2, 1, 6, 3, 4, 5, 7
    c. 2, 1, 7, 6, 5, 3, 4
    d. 7, 6, 1, 3, 4, 5, 2
22. What is/are true about the page table?
    a. It is an array indexed by virtual page numbers
    b. It resides in RAM
    c. It stores location in swap space on disk
    d. a and b
    e. All of the above
    f. None of the above

Name:_____ Student ID:_____ Section:____

23. Which mapping function (s) is/are generally used to implement virtual memory
    a. Fully associative placement
    b. Set-associative placement
    c. Direct mapping
    d. a and b
    e. All of the above
24. Which combination is not possible, when virtual memory is implemented in a computer system?
    a. A TLB miss, a page table hit, and a data cache hit
    b. A TLB miss, a page table hit, and a data cache miss
    c. A TLB miss, a page table miss, and a data cache miss
    d. A TLB miss, a page table miss, and a data cache hit
25. In a multi-processor environment, which protocol can be used to ensure cache coherency?
    a. Snooping protocol
    b. Write through protocol
    c. Directory-based protocol
    d. Cache update protocol
    e. a and c
    f. a, b, and c
    g. all of the above

Answer sheet for Part I

1	**9**	**17**
2	**10**	**18**
3	**11**	**19**
4	**12**	**20**
5	**13**	**21**
6	**14**	**22**
7	**15**	**23**
8	**16**	**24**
		25

Name:_____ Student ID:_____ Section:____

Part II (10 points)

For a 2-way associative cache design with a 32-bit address, the following bits of the address are used to access the cache.

Tag Index Offset

TAG	INDEX	OFFSET
31–9	8–5	4–0

a) What is the cache block size (in words)?

b) How many entries does the cache have in total?

c) Starting from power on, the following byte-addressed cache references are recorded.

Address	0	4	16	132	232	160	1024	30	140	3100	180	2180

If the LRU policy is used, List the state of the cache in the table below.

Address	0	4	16	132	232	160	1024	30	140	3100	180	2180
Index												
Set id												
Hit/Miss												

Name:_____ Student ID:_____ Section:____

Part III (40 points)

(a) (10 pts) Consider a computer and its disk storage system with the following specification:
Program instructions per I/O operation: **100,000**
OS instructions per I/O operation: **150,000**
Workload (KB): **64**
Processor speed (instructions per second): **4 Billion**
I/O bus bandwidth: **1,000 MB/s**
Disk average seek time: **6 ms** (applied to all cases)
Disk RPM: **7,200**
Disk transfer rate: **90 MB/s**
Number of disks supported by the controller: **8**
Find the maximum sustainable I/O rate for random disk reads/writes and the number of disks and controllers required to reach this rate. Which part of the system is a performance bottleneck in this scenario?

Name:_____ Student ID:_____ Section:____

(b) (10 pts) Describe RAID 0 and RAID 5 storage array configurations. In a hybrid RAID 5+0 configuration, sets of RAID 5 arrays are nested within a single RAID 0 array. Draw the RAID 5+0 storage array diagram and how the data and parity blocks are distributed.

Name:_____ Student ID:_____ Section:____

(c) (5 pts) Consider an output I/O transaction (i.e., data is read from a memory to an I/O device) in an asynchronous bus handshaking protocol. Describe what goes on with the ReadReq, Data, Ack, and DataRdy signals during the transaction.

(d) (5 pts) Discuss the advantages and disadvantages of polling versus Interrupt-driven I/O. For a computer controlling an aircraft autopilot system, which one is more suitable? Why?

Name:_____ Student ID:_____ Section:____

For question (e) and (f), consider two different processors, P1 and P2, implementing the same instruction set architecture. The CPIs for 4 instruction classes for each processor are listed in the table below.

Processor	Clock rate (GHz)	CPI class A	CPI clsss B	CPI class C	CPI class D
P1	1.8	1	2	3	3
P2	2.2	2	2	2	2

Given a program I_1 with 10^5 instructions divided into classes as follows: 15% class A, 10% class B, 50% class C, and 25% class D.

(e) (5 pts) Calculate the time spent executing program I_1 by each processor. Which one is faster?

(f) (5 pts) For each processor, determine the weighed average CPI and the number of clock cycles used to complete program I_1.

Name:_____ Student ID:_____ Section:____

Scratch note (กระดาษทด) -- you can tear this page off