

Kiểm tra Giữa kỳ 2009

Kiến trúc Máy tính Nâng cao

Thời gian: 90 phút

Câu 1. (3 điểm) Khảo sát sự khác nhau về miss penalty trong cache L1 giữa một cache L2 theo phương thức ánh xạ trực tiếp và một cache L2 theo phương thức 2-way set associative. Phương thức set associative làm tăng “hit time” 10% chu kỳ clock của CPU. “Hit time” của cache L2 theo phương thức ánh xạ trực tiếp là 10 chu kỳ clock. “Miss rate” cục bộ của cache L2 theo phương thức ánh xạ trực tiếp là 25% và theo phương thức 2-way set associative là 20%. “Miss penalty” của cache L2 là 50 chu kỳ clock. Do cache L2 được đồng bộ với cache L1 và clock của CPU, hit và miss phải được làm tròn thành một số nguyên chu kỳ xung clock.

	Direct cache	2-way cache
L2 hit time	10 cycles	$110\% \times 10 = 11 \text{ cycles}$
L2 miss rate	25%	20%
Miss penalty	50 cycles	50 cycles
AMAT	$75\% \times 10 + 25\% \times 50 = 20 \text{ cycles}$	$80\% \times 11 + 20\% \times 50 = 18.8 \approx 19 \text{ cycles}$

Câu 2. (4 điểm) Giả sử thời gian truy xuất bộ nhớ chính là 70ns và 36% truy xuất bộ nhớ này là các lệnh. Cache L1 được nối với 2 BXL P1 và P2 có các thông tin sau:

	L1 size	L1 miss rate	L1 hit time
P1	1 KB	11.4%	0.62ns
P2	2 KB	8%	0.66ns

- Giả sử hit time của L1 xác định thời gian chu kỳ cho các BXL. Xác định tốc độ của các BXL?
- Xác định AMAT cho các BXL
- Giả sử $CPI_{ideal} = 1$, xác định CPI thực tế của các BXL P1 và P2. BXL nào chạy nhanh hơn?

Giả sử thêm cache L2 với các đặc tính được cho trong bảng sau vào BXL P1. Miss rate của L2 cho trong bảng là miss rate cục bộ của riêng nó.

L2 size	L2 miss rate	L2 hit time
512 KB	98%	3.22ns

- Xác định lại AMAT của BXL P1 sau khi thêm cache L2. AMAT này tốt hơn hay kém hơn?
 - Giả sử $CPI_{ideal} = 1$, xác định CPI thực tế của P1
 - Với BXL P1 được thêm cache L2, P1 hay P2 nhanh hơn? Nếu P1 nhanh hơn, miss rate của cache L1 trong BXL P2 phải bằng bao nhiêu để hiệu suất của P2 bằng với hiệu suất P1? Nếu P2 nhanh hơn, miss rate của cache L1 trong BXL P1 phải bằng bao nhiêu để hiệu suất của P1 bằng với hiệu suất P2?
- a) $P1 = 1/0.62 \times 10^{-9} = 1.613 \text{ GHz}$
 $P2 = 1/0.66 \times 10^{-9} = 1.515 \text{ GHz}$
- b) Đối với P1: $AMAT = \text{hit time} + \text{miss rate} \times \text{miss penalty}$

- $$= (100\% - 11.4\%) \times 0.62 + 11.4\% \times 70 = 8.53 \text{ ns}$$
- Đối với P2: $AMAT = \text{hit time} + \text{miss rate} \times \text{miss penalty}$
- $$= (100\% - 8\%) \times 0.66 + 8\% \times 70 = 6.21 \text{ ns}$$
- c) P1: $CPI_{\text{real}} = CPI_{\text{ideal}} + \text{avg memory stall cycle} = 1 + 36\% \times 11.4\% \times 70 / 0.62 = 5.63$
P2: $CPI_{\text{real}} = CPI_{\text{ideal}} + \text{avg memory stall cycle} = 1 + 36\% \times 8\% \times 70 / 0.66 = 4.05$
CPU time P1: $= IC \times 5.63 \times 0.62 = 3.49 \times IC$
CPU time P2: $= IC \times 4.05 \times 0.66 = 2.67 \times IC$
 \Rightarrow P2 nhanh hơn
- d) Sau khi thêm cache L2 cho bộ xử lý P1:
 $AMAT = 0.62 + 11.4\% \times 3.22 + 11.4\% \times 98\% \times 70 = 8.81 \text{ ns}$
Rõ ràng AMAT này xấu hơn
- e) $CPI_{\text{real}} = CPI_{\text{ideal}} + \text{avg memory stall cycle} = 1 + 36\% \times 11.4\% \times 3.22 / 0.62 + 36\% \times 11.4\% \times 98\% \times 70 / 0.62 = 5.76$
- f) P1 có L2 cache: $CPI = 5.76$, trong khi P2: $CPI = 4.05$.
CPU time P1: $= IC \times 5.76 \times 0.62 = 3.57 \times IC$
CPU time P2: $= IC \times 4.05 \times 0.66 = 2.67 \times IC$
 \Rightarrow P2 vẫn nhanh hơn P1 mặc dù chỉ có cache L1
Muốn P1 nhanh như P2: $CPI_1 = 2.67/0.62 = 4.31$
 $1 + 36\% \times MR\% \times 3.22 / 0.62 + 36\% \times MR\% \times 98\% \times 70 / 0.62 = 4.31$
 $\Rightarrow MR = 7.93\%$

Câu 3. (1,5 điểm) Cho một bộ xử lý có 32 bit địa chỉ. Gọi S là tổng kích thước của bộ nhớ cache, B là kích thước một dòng trong cache và A là tổng số các khối dữ liệu có cùng chỉ mục mà cache có thể chứa trong cùng một thời điểm. Giả sử S, B và A đều là các giá trị lũy thừa của 2. Xác định các công thức sau theo S, B và A:

- Số bit cần thiết cho độ dời trong một khối (block offset)
 - Số tập hợp có trong cache
 - Số "ways" có trong cache
 - Số bit cần thiết cho việc chỉ mục (xác định dòng trong cache)
 - Số bit cần thiết cho tag
- a) Số bit cần thiết cho độ dời trong một khối: $w = \log_2 B$
b) Số khối có trong cache: S/B
 \Rightarrow Số tập hợp có trong cache: $v = (S/B)/A$
c) Số "ways" có trong cache: $k = S/B$
d) Số bit cần thiết cho việc chỉ mục: $\log_2(S/B)$
e) Số bit cần thiết cho tag: $32 - d - w = 32 - \log_2(S/B/A) - \log_2 B$

Câu 4. (1,5 điểm) Máy tính A được hiện thực theo kiến trúc tập lệnh MIPS và có tần số xung nhịp 2.5GHz. Máy tính B được hiện thực theo kiến trúc x86 và có tần số xung nhịp 3Ghz. Trung bình, các lệnh của chương trình MIPS có nhiều hơn 1.5 lần so với các lệnh của chương trình x86.

- Đối với chương trình P1, máy tính A có CPI là 2 và Máy tính B có CPI là 3. Máy tính nào chạy chương trình P1 nhanh hơn? Nhanh hơn bao nhiêu lần?
- Đối với chương trình P2, Máy tính A có CPI là 1 và Máy tính B có CPI là 2. Máy tính nào chạy chương trình P2 nhanh hơn? Nhanh hơn bao nhiêu lần?

- a) Giả sử chương trình P1 có n lệnh x86, do đó chương trình MIPS có $1.5n$ lệnh. Máy tính A có tần số xung clock là 2.5GHz, tức một chu kỳ có 0.4 ns. Thời gian thực hiện P1 trên máy A là $0.4 \text{ ns/clock} \times 2 \text{ clocks/instructions} \times 1.5n \text{ instructions} = 1.2n \text{ ns}$. Máy tính B hoạt động ở tần số 3GHz, do đó thời gian chương trình chạy trên máy B là $0.333 \times 3 \times n = n \text{ ns}$. Vậy máy tính B chạy nhanh hơn 1.2 lần.
- b) Tính toán tương tự, máy A thực hiện P2 trong $0.4 \times 1 \times 1.5n = 0.6n \text{ ns}$, trong khi máy B mất $0.333 \times 2 \times n = 0.666n \text{ ns}$. Vậy máy A nhanh hơn 1.11 lần.

Kiểm tra Giữa kỳ 2009

Kiến trúc Máy tính Nâng cao

Thời gian: 90 phút

Câu 1. (3 điểm) Khảo sát sự khác nhau về miss penalty trong cache L1 giữa một cache L2 theo phương thức ánh xạ trực tiếp và một cache L2 theo phương thức 2-way set associative. Phương thức set associative làm tăng “hit time” 10% chu kỳ clock của CPU. “Hit time” của cache L2 theo phương thức ánh xạ trực tiếp là 10 chu kỳ clock. “Miss rate” cục bộ của cache L2 theo phương thức ánh xạ trực tiếp là 25% và theo phương thức 2-way set associative là 20%. “Miss penalty” của cache L2 là 50 chu kỳ clock. Do cache L2 được đồng bộ với cache L1 và clock của CPU, hit và miss phải được làm tròn thành một số nguyên chu kỳ xung clock.

Câu 2. (4 điểm) Giả sử thời gian truy xuất bộ nhớ chính là 70ns và 36% truy xuất bộ nhớ này là các lệnh. Cache L1 được nối với 2 BXL P1 và P2 có các thông tin sau:

	L1 size	L1 miss rate	L1 hit time
P1	1 KB	11.4%	0.62ns
P2	2 KB	8%	0.66ns

- Giả sử hit time của L1 xác định thời gian chu kỳ cho các BXL. Xác định tốc độ của các BXL?
- Xác định AMAT cho các BXL
- Giả sử $CPI_{ideal} = 1$, xác định CPI thực tế của các BXL P1 và P2. BXL nào chạy nhanh hơn?

Giả sử thêm cache L2 với các đặc tính được cho trong bảng sau vào BXL P1. Miss rate của L2 cho trong bảng là miss rate cục bộ của riêng nó.

L2 size	L2 miss rate	L2 hit time
512 KB	98%	3.22ns

- Xác định lại AMAT của BXL P1 sau khi thêm cache L2. AMAT này tốt hơn hay kém hơn?
- Giả sử $CPI_{ideal} = 1$, xác định CPI thực tế của P1
- Với BXL P1 được thêm cache L2, P1 hay P2 nhanh hơn? Nếu P1 nhanh hơn, miss rate của cache L1 trong BXL P2 phải bằng bao nhiêu để hiệu suất của P2 bằng với hiệu suất P1? Nếu P2 nhanh hơn, miss rate của cache L1 trong BXL P1 phải bằng bao nhiêu để hiệu suất của P1 bằng với hiệu suất P2?

Câu 3. (1,5 điểm) Cho một bộ xử lý có 32 bit địa chỉ. Gọi S là tổng kích thước của bộ nhớ cache, B là kích thước một dòng trong cache và A là tổng số các khối dữ liệu có cùng chỉ mục mà cache có thể chứa trong cùng một thời điểm. Giả sử S, B và A đều là các giá trị lũy thừa của 2. Xác định các công thức sau theo S, B và A:

- Số bit cần thiết cho độ dời trong một khối (block offset)
- Số tập hợp có trong cache
- Số "ways" có trong cache
- Số bit cần thiết cho việc chỉ mục (xác định dòng trong cache)
- Số bit cần thiết cho tag

Câu 4. (1.5 điểm) Giả sử một chương trình có tỉ lệ các lệnh như sau: 30% lệnh load, 10% lệnh store, 50% lệnh add, 8% lệnh multiply, và 2% lệnh divide. Giả sử bộ xử lý có thời gian thực hiện các lệnh như sau: loads: 4 cycles, stores: 4 cycles, adds: 2 cycles, multiplies: 16 cycles, và divides: 50 cycles. Nếu phải cải tiến bộ xử lý này để nó chạy nhanh hơn, lệnh nào trong các loại lệnh trên sẽ được ưu tiên tối ưu trước? Tại sao?

Số chu kỳ của các lệnh là

Lệnh	%	CPI	Cycles
load	30	4	120 cycles
store	10	4	40 cycles
add	50	2	100 cycles
multiply	8	16	128 cycles
divide	2	50	100 cycles

Lệnh Multiply chiếm nhiều chu kỳ nhất, nên sẽ được ưu tiên tối ưu trước.