

Trường Đại Học Bách Khoa Tp. Hồ Chí Minh
Khoa Khoa Học & Kỹ Thuật Máy Tính



MÔN HỌC KIẾN TRÚC MÁY TÍNH NÂNG CAO

Bài tập lớn 3

Cache

CBGD:
TS. Trần Ngọc Thịnh

Nhóm 1:
Hồ Quang Chi Bảo - 7140219
Phạm Thế Anh - 7140217
Trần Nhật Hoàng Anh - 7140218
Nguyễn Hoàng Biết - 7140220
Lê Trương Trọng Duy - 7140227
Lê Thanh Bình - 7140221



Mục lục

1	Đề bài	2
2	Bài giải	3
2.1	Bài 1	3
2.2	Bài 2	4
2.3	Bài 3	5



1 Đề bài

Bài 1.

Design an L1 cache (number of bits for tag, entry, ...) for a CPU with 32-bit address in 3 following types. The cache size is 32KB, block (line) size is 32 byte

- Direct mapped
- Fully Associative
- 4-way Associative

Bài 2.

For the cache in Exercise 1, assumed the cache is 4-way associative, how many cache hit, miss occur if the CPU execute the following memory access sequence in case of

- Write through no write allocation
- Write back with write allocation

1	RD 0x00000000
2	WR 0x01000000
3	RD 0x01000010
4	WR 0x02000050
5	RD 0x02000058

Bài 3.

How much faster/slower is a unified 32KB cache than a separated 16KB I/16KB D cache if the miss rate is ones in the following slide table, and there are 70% instructions are LD/ST. Assumed that unified cache has 1-port only. The hit time is 1 cycle and miss penalty is 50 cycles



2 Bài giải

2.1 Bài 1

Theo đề bài ta có:

- Address bits $N = 32$
- Block size = $32B$. Ta tính được bit offset b như sau: $\text{Block size} = 32 \Leftrightarrow 2^b = 32 \Leftrightarrow b = 5 \text{ bits}$
- Cache size = $32KB \Rightarrow$ số lượng cache line: $\text{Number of cache line} = 32KB/32B = 1024 = 2^{10} \text{ lines}$

Ta tính số index bit và tag bit cho từng loại cache như sau:

- Direct mapped cache:
 - Index bit k : $2^k = 2^{10} \Rightarrow k = 10 \text{ bits}$
 - Tag bit t : $t = N - k - b = 32 - 10 - 5 = 17 \text{ bits}$
- Fully associative cache
 - Index bit $k = 0$
 - Tag bit t : $t = N - k - b = 32 - 0 - 5 = 27 \text{ bits}$
- 4-way set associative cache
 - Vì là sử dụng loại 4-way set associative cache nên ta có $2^w = 4 \Rightarrow w = 2$
 - Số lượng set = $1024/4 = 256 \Rightarrow$ Index bit $k = 8 \text{ bits}$
 - Tag bit t : $t = N - k - b = 32 - 8 - 5 = 19 \text{ bits}$



2.2 Bài 2

Tính số lần cache hit và cache miss khi CPU thực hiện đoạn lệnh đã cho khi sử dụng 4-way set associative cache cho từng cơ chế write cache. Khi khởi động, toàn bộ cache là empty.

- Write through no write allocation: tổng cộng có 5 lần cache miss như sau:
 - RD 0x00000000 → cache miss, load 1 block 32 bytes from 0x00000000 to 0x0000001C to way0 of set 0
 - WR 0x01000000 → cache miss, no allocate
 - RD 0x01000010 → cache miss, load 1 block 32 bytes from 0x01000000 to 0x0100001C to way1 of set 0
 - WR 0x02000050 → cache miss, no allocate
 - RD 0x02000058 → cache miss, load 1 block 32 bytes from 0x02000040 to 0x0200005C to way2 of set 0
- Write back with write allocation: tổng cộng có 3 lần cache miss và 2 lần cache hit như sau:
 - RD 0x00000000 → cache miss, load 1 block 32 bytes from 0x00000000 to 0x0000001C to way0 of set 0
 - WR 0x01000000 → cache miss, load 1 block 32 bytes from 0x01000000 to 0x0100001C to way1 of set 0
 - RD 0x01000010 → cache hit, previous load including this address
 - WR 0x02000050 → cache miss, load 1 block 32 bytes from 0x02000040 to 0x0200005C to way2 of set 0
 - RD 0x02000058 → cache hit, previous load including this address



2.3 Bài 3

Để so sánh sự nhanh/chậm hơn giữa 2 loại cache theo đề bài, ta tính thời gian truy cập bộ nhớ trung bình $AMAT$ của từng loại. Ta có công thức tính $AMAT$ tổng quát như sau:

$$AMAT = \%Instructions * AMAT_{Instructions} + \%Data * AMAT_{Data} \quad (1)$$

Trong đó, $AMAT$ của lệnh và dữ liệu được tính bằng công thức:

$$AMAT = HitTime + MissRate * MissPenalty \quad (2)$$

Công thức tính Miss Rate được tính từ bảng đã cho như sau:

$$MissRate = \frac{Misses\ per\ 1000\ instructions / 1000}{Memory\ Accesses / Instruction} \quad (3)$$

Theo đề bài, ta tính được:

$$\%Instructions = \frac{100\%}{100\% + 70\%} = 0.588 \quad (4)$$

$$\%Data = \frac{70\%}{100\% + 70\%} = 0.412 \quad (5)$$

Thay thế số vào ta được:

- Unified 32KB cache

$$MissRate_{32kb\ Unified} = \frac{43.3/1000}{1 + 0.7} = 0.0255$$

Giả sử Unified cache là loại single port, do đó không thể truy cập đồng thời lệnh và dữ liệu trong một chu kỳ \Rightarrow CPU sẽ bị stall thêm 1 chu kỳ. Vậy ta tính được $AMAT$ cho loại cache này như sau:

$$\begin{aligned} AMAT_{32kb\ Unified} &= 0.588 * (1 + 0.0255 * 50) + 0.412 * (1 + 1 + 0.0255 * 50) \\ &= 2.68 \text{ (clock cycles)} \end{aligned}$$

- Separated 16KB I/16KB D cache

$$\begin{aligned} AMAT_{16kb\ instruction} &= 1 + \frac{3.82/1000}{1} * 50 = 1.191 \\ AMAT_{16kb\ data} &= 1 + \frac{40.9/1000}{0.7} * 50 = 3.921 \\ AMAT_{overall\ Separated} &= 0.588 * 1.191 + 0.412 * 3.921 \\ &= 2.31 \text{ (clock cycles)} \end{aligned}$$

Vậy nếu ta dùng Separated 16KB I/16KB D cache sẽ nhanh hơn $2.68/2.31 = 1.16$ lần so với Unified 32KB cache.