**Bộ Nhớ**

**Tổng Quan Về Hệ Thống Nhớ**

**Các đặc trưng của hệ thống nhớ**

Bên trong CPU:

* tập thanh ghi

Bộ nhớ trong

* bộ nhớ chính
* bộ nhớ cache

Bộ nhớ ngoài: các thiết bị nhớ

**Phương pháp truy nhập**

* Tuần tự (bằng từ)
* Trực tiếp (các loại đĩa)
* Ngẫu nhiên (bộ nhớ bán dẫn)
* Liên kết (cache)

**Hiệu năng**

* Thời gian truy nhập
* Chu kỳ nhớ
* Tốc độ truyền

**Kiểu vật lý**

* Bán dẫn
* Từ
* Quang

**Đặc tính vật lý**

* Khả biến / không khả biến
* Xóa được / không xóa được
* Tổ chức

**Phân Cấp Hệ Thống Nhớ**

**Chiều từ trên xuống dưới:**

* Dung lượng tăng dần
* Tốc độ giảm dần
* Giá thành/bit giảm dần

**Bộ Nhớ Bán Dẫn**

**ROM (Read Only Memory)**

* Bộ nhớ không khả biến
* Lưu trữ: thư viện chương trình con, BIOS, bảng chức năng, vi chương trình
* Không thể ghi lại sau khi sản xuất
* Các kiểu:
  + ROM mặt nạ (Mask ROM)
    - Được lập trình trong quá trình sản xuất, rất đắt
  + PROM (Programmable ROM)
    - Cần thiết bị chuyên dụng để ghi bằng chương trình → chỉ ghi được một lần
  + EPROM (Erasable Programmable ROM)
    - Có thể xóa và lập trình lại bằng tia UV
  + EEPROM (Electrically Erasable Programmable ROM)
    - Có thể xóa và lập trình lại bằng điện
  + Flash memory
    - ghi theo khối, xóa bằng điện

**RAM (Random Access Memory)**

* Bộ nhớ đọc-ghi, khả biến, lưu thông tin tạm thời
* **Phân loại:**
  + **SRAM (Static RAM)**
    - Dùng Flip-Flop
    - cấu trúc phức tạp, đắt
    - Tốc độ nhanh, thường làm cache
  + **DRAM (Dynamic RAM)**
    - Dùng tụ điện
    - Cần làm tươi, chậm hơn
    - Các DRAM tiên tiến:
      * EDRAM (Enhanced DRAM)
      * CDRAM (Cache DRAM)
      * SDRAM
      * DDR-SDRAM
      * RDRAM

**Bộ Nhớ Chính**

* Chứa chương trình và dữ liệu đang xử lý
* Tồn tại trên mọi hệ thống
* CPU đánh địa chỉ trực tiếp
* Dung lượng thường nhỏ hơn không gian địa chỉ của CPU
* **Tổ chức bộ nhớ đan xen (Interleaved)**
* Độ rộng bus dữ liệu: 8/16/32-bit
* Tổ chức vật lý tùy thuộc số byte

**Bộ Nhớ Đệm Nhanh (Cache)**

**Đặc điểm**

* Dựa trên nguyên lý cục bộ hóa tham chiếu
* Cache nằm giữa CPU và bộ nhớ chính để tăng tốc
* CPU truy vấn Cache trước

**Cấu trúc**

* Bộ nhớ chính = 2^N byte → chia thành Block
* Cache = tập hợp các Line
* Mỗi Line mang thông tin Tag
* Cache hit / Cache miss

**Các Phương Pháp Ánh Xạ**

**Phương pháp ánh xạ**: là các phương pháp tổ chức bộ nhớ cache.

* Ánh xạ trực tiếp (Direct mapping)
* Ánh xạ liên kết toàn phần (Fully associative mapping)
* Ánh xạ liên kết tập hợp (Set associative mapping)

**1. Ánh xạ trực tiếp**

* Mỗi Block chỉ nạp vào **một Line** duy nhất
* Địa chỉ gồm:
  + Word
  + Line
  + Tag

**2. Ánh xạ liên kết toàn phần**

* Mỗi Block có thể nạp vào **bất kỳ Line nào** trong cache
* Địa chỉ gồm:
  + Word
  + Tag

**3. Ánh xạ liên kết tập hợp**

* Cache chia thành các **Set**, mỗi Set có 1 số Line
* Địa chỉ gồm:
  + Word
  + Set
  + Tag

**Ví Dụ Về Ánh Xạ Địa Chỉ**

**Ví dụ:**

* Không gian địa chỉ = 4GB
* Cache = 256KB
* Block = 32 byte

**Số bit các trường:**

* Ánh xạ trực tiếp: Tag = 14 bit
* Ánh xạ liên kết toàn phần: Tag = 27 bit
* Ánh xạ liên kết tập hợp (4 đường): Tag = 16 bit

**Thuật Giải Thay Thế**

* **Ánh xạ trực tiếp:** Thay thế Block ở chính Line đó
* **Ánh xạ liên kết:** Các thuật toán:
  + Random: thay thế ngẫu nhiên
  + FIFO: thay thế Block cũ nhất
  + LFU: (Least Frequently Used): thay thế Block ít được sử dụng nhất trong một khoảng thời gian
  + LRU (Least Recently Used): thay thế Block ít được sử dụng nhất

**Phương Pháp Ghi Dữ Liệu Khi Cache Hit**

* **Ghi xuyên qua (Write-through):** Ghi cả cache + bộ nhớ chính (chậm)
* **Ghi trả sau (Write-back):** Chỉ ghi cache → chỉ ghi bộ nhớ chính khi Block trong cache bị thay thế

**Cache Trên Các Bộ Xử Lý Intel**

| **Bộ xử lý** | **Cache** |
| --- | --- |
| 80486 | 8KB L1 trên chip |
| Pentium | 2 cache L1 trên chip (8KB lệnh + 8KB dữ liệu) |
| Pentium 4 (2000) | 2 mức cache L1 và L2 trên chip |

**Củng Cố**

* Các loại bộ nhớ
* Phân cấp bộ nhớ
* Bộ nhớ bán dẫn
* Bộ nhớ cache
* Các phương pháp ánh xạ
* Các thuật giải thay thế