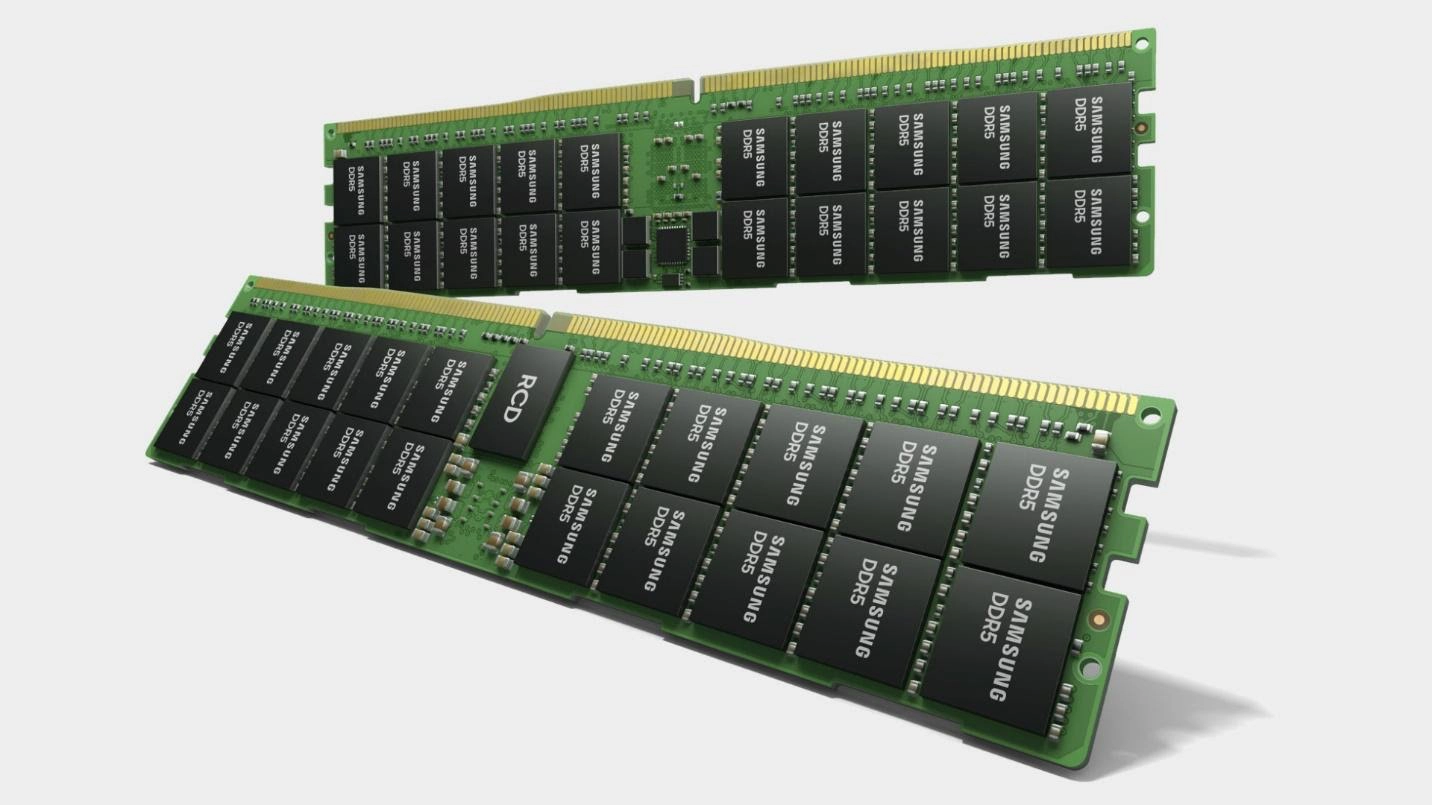
**BỘ NHỚ RAM, CACHE**

**VÀ**

**BỘ NHỚ CHÍNH**



**① Bộ nhớ trong máy tính.**

### Bộ nhớ máy tính là gì ?

**Bộ nhớ máy tính** (tiếng anh: Computer data storage), thường được gọi là ổ nhớ (storage) hoặc bộ nhớ (memory), là một thiết bị công nghệ bao gồm các phần tử máy tính và lưu trữ dữ liệu, được dùng để duy trì dữ liệu số. Nó là một linh kiện cơ bản có chức năng cốt lõi của các máy tính.

Trên thực tế, bộ nhớ máy tính cũng được chia ra làm nhiều loại có những tính năng và công dụng khác nhau bao gồm:

**Bộ nhớ ngoài:**

Tất cả các thiết bị lưu trữ thứ cấp có khả năng lưu trữ dữ liệu khối lượng lớn được gọi là bộ nhớ ngoài, bộ nhớ phụ. Nó chậm hơn bộ nhớ chính. Tuy nhiên, nó có thể tiết kiệm một lượng dữ liệu đáng kể, trong phạm vi từ gigabyte đến terabyte. Bộ nhớ này còn được gọi là bộ nhớ ngoài hoặc phương tiện lưu trữ đại chúng.

Một số loại / linh kiện bộ nhớ ngoài phổ biến:

* Đĩa mềm.



* Đĩa cứng (HDD, SSD).



* USB / Flash.



* Đĩa quang



* Băng từ



**Bộ nhớ trong.**

Bộ nhớ tronghay Internal Memory chính là khái niệm chỉ những loại bộ nhớ đã được lắp đặt sẵn cũng như sử dụng trong những thiết bị ví dụ như máy tính bảng, điện thoại hoặc máy tính. Thông thường bộ nhớ trong máy tính sẽ có hai loại chính đó chính là bộ nhớ đệm (Cache) và bộ nhớ chính (ROM, RAM).

Bộ nhớ trong thuộc loại bộ nhớ có thể truy cập dữ liệu ngẫu nhiên.

**②Bộ nhớ cache.**

### Bộ nhớ cache là gì ?

Bộ nhớ Cache chính là một trong những thành phần thuộc bộ nhớ trong. Theo đó nó có tác dụng lưu trữ những thông tin, dữ liệu được dùng thường xuyên để giúp cho CPU có thể truy cập trong tương lai với tốc độ nhanh chóng hơn. Bộ nhớ đệm nhìn chung đã nằm sẵn ở trong máy tính, đồng thời cũng có tác dụng tương tự như những thành  RAM được cắm trên mainboard.

Bộ nhớ cache là một loại bộ nhớ tạm thời nằm giữa bộ nhớ trung tâm và bộ xử lý trong máy tính. Nó được sử dụng để lưu trữ dữ liệu và thông tin truy cập gần đây nhất của CPU, giúp tăng tốc độ xử lý dữ liệu và giảm thời gian truy cập đến bộ nhớ chính. Bộ nhớ cache thường được tích hợp trực tiếp vào chip xử lý hoặc được đặt riêng biệt trên bo mạch chủ. Có nhiều loại bộ nhớ cache khác nhau, bao gồm cache L1, L2 và L3, với kích thước và tốc độ truy xuất khác nhau.

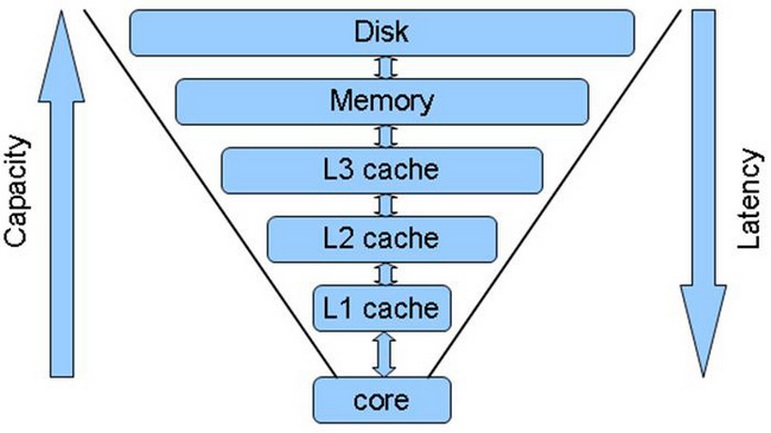
Trong bộ nhớ đệm cache nếu như càng lớn thì khi đó dung lượng sẽ càng lớn, hoạt động mượt hơn và cũng có nhiều không gian để lưu trữ hơn.

Cấu tạo của bộ nhớ cache bao gồm:

1. Bộ nhớ cache L1: Là loại bộ nhớ cache nhỏ nhất và nhanh nhất, được tích hợp trực tiếp vào chip xử lý. Bộ nhớ L1 thường có dung lượng nhỏ, từ 8KB đến 64KB, và được chia thành hai phần: cache lệnh (instruction cache) và cache dữ liệu (data cache). Cache lệnh lưu trữ các lệnh được CPU thực hiện, còn cache dữ liệu lưu trữ các dữ liệu được xử lý bởi CPU.

2. Bộ nhớ cache L2: Là loại bộ nhớ cache lớn hơn L1, được đặt riêng biệt trên bo mạch chủ. Bộ nhớ L2 có dung lượng lớn hơn, từ 512KB đến 8MB, và được chia thành hai phần tương tự như L1. Tuy nhiên, bộ nhớ L2 có tốc độ truy xuất chậm hơn L1.

3. Bộ nhớ cache L3: Là loại bộ nhớ cache lớn nhất, được đặt trên bo mạch chủ và dùng chung cho tất cả các lõi xử lý của CPU. Bộ nhớ L3 có dung lượng lớn, từ 4MB đến 64MB, và có tốc độ truy xuất chậm hơn L2.



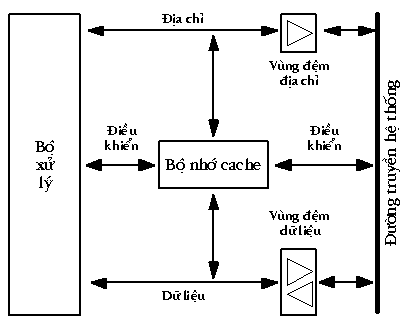
*Vị trí của bộ nhớ cache trong máy tính*

Bộ nhớ cache được chia thành nhiều khối nhỏ hơn gọi là cache line hoặc block. Kích thước của cache line thường là 64 byte hoặc 128 byte. Mỗi cache line chứa một phần của dữ liệu được lưu trữ trong bộ nhớ chính. Khi CPU yêu cầu truy cập dữ liệu, bộ nhớ cache sẽ kiểm tra xem dữ liệu đó đã được lưu trữ trong cache hay chưa. Nếu đã có trong cache, CPU sẽ lấy dữ liệu từ cache line thay vì phải truy cập đến bộ nhớ chính, giúp tăng tốc độ xử lý. Nếu dữ liệu chưa có trong cache, CPU sẽ phải truy cập đến bộ nhớ chính để lấy dữ liệu và lưu trữ vào cache để sử dụng cho những lần truy cập sau. Việc lưu trữ dữ liệu trong cache được quản lý bằng cách sử dụng các thuật toán thay thế (replacement algorithms) như LRU (Least Recently Used) hoặc FIFO (First In First Out).

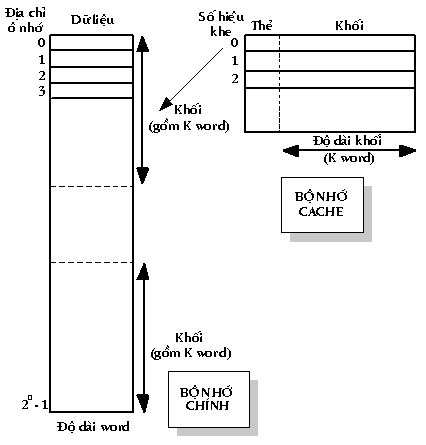
Bộ nhớ cache được thiết kế với ý định mang lại bộ nhớ có tốc độ xấp xỉ tốc độ của bộ nhớ nhanh nhất hiện có, đồng thời cung cấp một kích thước bộ nhớ lớn với phí tổn ít hơn so với các loại bộ nhớ bán dẫn. Khái niệm về bộ nhớ cache được trình bày trên hình 2.1 và 2.2. Bộ nhớ cache chứa bản sao của một phần bộ nhớ chính. Khi CPU cố gắng đọc một word từ bộ nhớ, word này sẽ được kiểm tra xem có trong cache hay không. Nếu có, word đó sẽ được cung cấp ngay cho CPU. Trong trường hợp ngược lại, một khối bộ nhớ chính, bao gồm một lượng cố định các word sẽ được đọc vào trong cache và sau đó word đó sẽ được cung cấp cho CPU. Do có hiện tượng tham chiếu cục bộ, khi một khối dữ liệu được lấy vào trong cache để đáp ứng một tham chiếu đơn lẻ, có nhiều khả năng các tham chiếu kế tiếp sẽ là các word khác trong cùng một khối.



*Hình 2.1 : Cache và bộ nhớ chính*



*Hình 2.2 : Tổ chức tiêu biểu của bộ nhớ cache*



*Hình 2.3 : Cấu trúc cache / bộ nhớ chính*

Hình 2.3 mô tả một tổ chức tiêu biểu của bộ nhớ cache. Bộ nhớ chính gồm đến 2n word khả định địa chỉ, với mỗi word có một địa chỉ duy nhất n bit. Do yêu cầu ánh xạ, bộ nhớ này được xem như gồm một số các khối có độ dài cố định K word. Tức là chúng ta có M = 2n/K khối. Bộ nhớ cache gồm C khe K word, và số các khe, hay đường, nhỏ hơn đáng kể so với số khối bộ nhớ chính (C << M). Tại một thời điểm bất kỳ, luôn có một tập con gồm các khối nhớ thường trú trong các khe của cache. Nếu một word trong một khối bộ nhớ được đọc, khối đó sẽ được chuyển vào một trong các khe của cache. Do số khối nhiều hơn số khe, một khe có thể không được giành cho một khối trong thời gian lâu dài. Vì lý do đó, mỗi khe có một thẻ cho biết khối nào đang được lưu trữ. Thẻ thường là một phần của địa chỉ bộ nhớ chính của khối đang được giữ trong khe.

Bộ nhớ đệm sẽ giúp cho máy tính xử lý nhanh hơn tuy nhiên nếu như để lâu và không xóa đi sẽ khiến cho lượng file rác tăng lên và giảm hiệu suất của máy tính xuống. Tuy rằng không nên quá thường xuyên làm điều này nhưng thỉnh thoảng cũng phải dọn dẹp nếu như cảm thấy cần thiết.

### Cách vận hành của bộ nhớ cache.

Phần cứng cài đặt cache như một nơi chứa dữ liệu tạm thời để có thể sử dụng lại. Vi xử lý (CPUs) và ổ đĩa cứng (HDD) thường xuyên sử dụng Cache,tương tự trình duyệt web và máy chủ (server).

Cache gồm thanh ghi. Mỗi thanh ghi chứa 1 bit để đánh dấu là đang lưu dữ liệu từ bộ nhớ hoặc chưa được sử dung, 1 nhãn (tag) để kiểm tra xem nó có phải là dữ liệu ứng với bộ nhớ hay không và 1 vùng để lưu trữ dữ liệu giống với bộ nhớ.

Khi cache client (vi xử lý,trình duyệt web,...) cần truy cập đến dữ liệu ở bộ nhớ,nó sẽ kiểm tra cache. Nếu 1 thanh ghi có nhãn tương ứng với dữ liệu mong muốn thì cache client sẽ sử dụng dữ liệu trên thanh ghi đó.Tình huống này được gọi là cache-hit. Ngược lại, nếu không tìm được thanh ghi tương ứng, thì tình huống này gọi là cache miss. Lúc này CPU sẽ truy cập vào bộ nhớ, lấy dữ liệu cần truy cập,đồng thời lưu dữ liệu đó vào 1 thanh ghi nào đó.

Trong trường hợp cache miss, CPU sẽ loại bỏ một số thanh ghi để dọn chỗ cho dữ liêu không nằm trong cache. Việc loại bỏ thanh ghi phải tuân thủ các quy định về thay thế,thông dụng nhất là "least recently used" (LRU), tức là loại bỏ những cache ít được sử dụng nhất.

### Nguyên tắc ghi đọc dữ liệu của hệ thống với bộ nhớ cache.

Khi CPU ghi dữ liệu vào Cache, cần phải có một số quy tắc ghi dữ liệu đó vào bộ nhớ.

Có 2 cách tiếp cận:

- Write-through: ghi ngay lập tức dữ liệu lên cả cache và bộ nhớ(DRAM)

-Write-back: ghi cho cache trước, việc ghi lên bộ nhớ bị trì hoãn đến khi có dữ liệu mới cần đặt lên vùng lưu trữ đó.

Một write-back cache được cài đặt khá phức tạp. Nó đánh dấu những thanh ghi có dữ liệu thay đổi,đánh dấu chúng là có thể ghi đè ("dirty"), khi thanh ghi bị thay thế, dữ liệu trên(được đánh dấu) sẽ được ghi vào bộ nhớ. Do đó,khi read-miss trên write-back cache,mất 2 lần truy cập,một là để ghi dữ liệu bị thay đổi vào bộ nhớ, hai là để lấy dữ liệu cần dùng và ghi lên cache.

Khi write-miss (xảy ra cache-miss), có 2 cách xử lý:

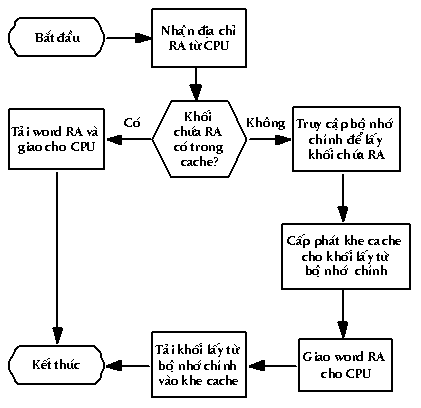
-Write allocate: dữ liệu sẽ được ghi lên cache, một thanh ghi sẽ bị thay thế. Cách này tương tự read-miss

-No-write allocate: dữ liệu sẽ được ghi vào bộ nhớ nhưng không ghi trên cache

Cả write-through và write-back đều sử dụng 2 cách trên khi write-miss, nhưng phổ biến chúng được ứng với từng cách:

-Write-back cache sử dụng write allocate giúp ghi đè dữ liệu trên cache hiệu quả hơn

-Write-through cache sử dụng no-write allocate và việc ghi dư liệu không được lưu trữ trong cache không có hiệu quả.



*Hình 2.4 minh họa thao tác đọc. CPU sinh ra địa chỉ, RA, của một word cần đọc. Nếu word có trong cache, nó sẽ được giao cho CPU. Ngược lại, khối chứa word sẽ được tải vào trong cache, và word được giao cho CPU.*

### Ứng dụng của bộ nhớ cache.

Bộ nhớ đệm CPU

Bộ nhớ đệm được sử dụng bởi đơn vị xử lý trung tâm (CPU) của máy tính để giảm thiểu thời gian trung bình trong việc truy cập dữ liệu trong bộ nhớ chính. Bộ nhớ đệm nhỏ hơn, nhanh hơn trong việc lưu trữ dữ liệu sao chép từ vùng bộ nhớ chính được sử dụng thường xuyên nhất. Hầu hết CPU sở hữu những bộ nhớ đệm độc lập khác biệt nhau, bao gồm lệnh và bộ nhớ đệm dữ liệu, nơi mà bộ nhớ đệm dữ liệu thường được thiết lập như một hệ thống cấp bậc của nhiều hơn các cấp bộ nhớ đệm.

Bộ nhớ đệm của vi xử lý đồ họa

Các phiên bản trước của vi xử lý đồ họa (GPUs) không có (các) bộ nhớ đệm cứng; tuy nhiên, khi GPUs chuyển sang vi xử lý không mang tính đồ họa, chúng lại sử dụng các bộ nhớ đệm lớn. Chẳng hạn như GT200 không có bộ đệm L2, trong khi bộ đệm của Fermi lại có 768KB, của Kepler là 1536KB và Maxwell có 2048KB.

Bộ đệm chuyển đổi

Đơn vị quản lý dữ liệu (MMU): tìm nạp những lối vào của trang bảng từ bộ nhớ chính sở hữu bộ nhớ đệm đặc biệt, được sử dụng để ghi lại kết quả của địa chỉ ảo đến địa chỉ vật lý. Bộ nhớ đệm đặc biệt này được gọi là Bộ Đệm Chuyển Đổi

Sự khác biệt giữa buffer và bộ nhớ cache

Hai từ ngữ "buffer" và "cache" không bao gồm nhau; thậm chí, có sự khác biệt cơ bản trong ý nghĩa giữa các quá trình của bộ nhớ cache và các quá trình của buffer.

Về cơ bản, bộ nhớ đệm nhận ra một sự gia tăng hiệu suất cho việc lấy đi lấy lại dữ liệu nhiều lần. Trong khi cache có thể nhận ra một sự gia tăng hiệu suất ngay từ ban đầu của một dữ liệu, quá trình tăng hiệu suất xảy ra trong hệ thống bộ nhớ đệm.

Với đọc cache, một mục dữ liệu phải được lấy từ vị trí gốc của nó ít nhất một lần để cho sau này đọc liên tiếp dữ liệu nhiều lần để thấy một sự gia tăng hiệu suất bằng hiệu lực của việc có thể lưu trữ trung gian của bộ nhớ cache nhanh hơn của dữ liệu gốc. Với ghi cache, một sự gia tăng hiệu suất của việc viết một dữ liệu có thể được nhận ra khi viết lần đầu tiên của các dữ liệu bằng hiệu lực của các dữ liệu được lưu trữ ngay trên bộ nhớ cache, trì hoãn việc chuyển giao các dữ liệu về sau hay nói cách khác nó xảy ra như là một tiến trình nền. Trái với đệm nghiêm ngặt, một quá trình trong bộ nhớ đệm phải tuân thủ một (phân phối tiềm năng) bộ nhớ cache giao thức mạch lạc để duy trì sự thống nhất giữa lưu trữ trung gian của bộ nhớ cache và vị trí nơi chứa dữ liệu. Buffering, mặt khác,

+ làm giảm số lượng các quá trình chuyển dữ liệu giữa quá trình giao tiếp, mà tối ưu trên không đề cập đến những truy cập một dữ liệu ít lần.

+ cung cấp quá trình giao tiếp trung gian

+ đảm bảo một kích thước dữ liệu tối thiểu hoặc đại diện theo yêu cầu của ít nhất một trong các quá trình giao tiếp liên quan đến việc truy cập.

Buffer là một vị trí bộ nhớ tạm thời thường được sử dụng vì các CPU instruction không thể trực tiếp truy cập các dữ liệu được lưu trữ trong các thiết bị ngoại vi. Như vậy, bộ nhớ địa chỉ được sử dụng như là một giai đoạn trung gian. Ngoài ra, một buffer như vậy có thể khả thi khi một khối lượng lớn các dữ liệu được lắp ráp hoặc tháo rời (theo yêu cầu của một thiết bị lưu trữ), hoặc khi dữ liệu có thể được gửi theo một thứ tự khác nhau hơn là trong đó nó được tạo ra. Ngoài ra, một bộ đệm toàn bộ dữ liệu thường được chuyển tuần tự (ví dụ vào đĩa cứng), vì vậy chính buffer đôi khi làm tăng hiệu suất truyền tải hoặc làm giảm sự biến đổi hoặc sự rối loạn của độ trễ việc chuyển giao như trái ngược với bộ nhớ đệm mà mục đích là để giảm độ trễ. Những lợi ích này có mặt ngay cả khi dữ liệu buffer được ghi vào buffer một lần và đọc từ bộ đệm một lần.

Một bộ nhớ cache cũng làm tăng hiệu suất truyền tải. Một phần của sự gia tăng tương tự xuất phát từ khả năng rằng nhiều giao dịch chuyển nhỏ sẽ kết hợp thành một khối lớn. Nhưng chính hiệu-được xảy ra bởi vì có một cơ hội tốt mà cùng một dữ liệu sẽ được đọc từ bộ nhớ cache nhiều lần, hoặc số liệu bằng văn bản sẽ sớm được đọc. Mục đích duy nhất của một bộ nhớ cache là để giảm truy cập vào kho chậm hơn nằm bên dưới. Bộ nhớ cache cũng thường là một lớp trừu tượng được thiết kế để có thể nhìn thấy từ quan điểm của các lớp lân cận.

Bộ nhớ đệm đĩa (Disk Cache)

Bộ nhớ đệm đĩa là một bộ phận để cải thiện thời gian đọc từ đĩa cứng hoặc ghi lên đĩa cứng. Hiện nay, bộ nhớ đệm đĩa thường được xem như một phần của đĩa cứng. Bộ nhớ đệm đĩa cũng có thể là một phần riêng biệt của RAM. Nó chứa dữ liệu đã được đọc gần đây và những vùng dữ liệu gần kề có thể được kết nối trong thời gian tới.Việc viết bộ nhớ đệm cũng được hộ trỡ bởi một số bộ nhớ đệm đĩa.

Một số bộ nhớ đệm khác

Những cỗ máy tìm kiếm (như Google, Yahoo,...) thường dùng bộ nhớ đệm để lưu giữ những trang web đã được tìm kiếm. Ví dụ, Google sẽ cung cấp một đường link bên cạnh kết quả tìm kiếm trỏ đến bộ nhớ đệm đối với những trang web đã được tìm kiếm trước đó. Điều này trở nên hữu ích khi những trang web từ một máy chủ nào đó trở nên khó truy cập trong thời gian ngắn hoặc lâu dài.

Một loại bộ nhớ đệm khác sẽ được dùng để lưu giữ những kết quả tính toán mà có thể sẽ được sử dụng trở lại. Ví dụ, cache là một chương trình tạo ra bộ nhớ đệm cho đầu ra của trình biên dịch để tăng tốc độ biên dịch cho những lần biên dịch sau.

Việc tạo bộ nhớ đệm cho cơ sở dữ liệu có thể cải thiện đáng kể khả năng vận hành của những ứng dụng cơ sở dữ liệu, như trong vận hành các đề mục, dữ liệu từ điển, hay những dạng thông dụng khác của dữ liệu.

## ③ Bộ nhớ chính

Bộ nhớ chính của máy vi tính dùng để chứa các thông tin cần thiết như chương trình, dữ liệu trong quá trình máy hoạt động. ROM và RAM là bộ nhớ chính của máy tính, dùng lưu trữ các chương trình quản lý việc khởi động (ROM) và các chương trình đang hoạt động trên máy tính (RAM). Ngày nay với công nghệ và kỹ thuật phát triển ROM và RAM được tạo ra với nhiều chủng loại khác nhau, đáp ứng mọi nhu cầu sử dụng của người dùng.

## ④ Bộ nhớ ROM

Một loại bộ nhớ với chức năng đọc, được nhà sản xuất ghi sẵn và chứa những chương trình giúp máy tinh có thể dễ dàng khởi động. ROM có chứa thông tin bảo mật như BIOS, bo mạch chủ máy tính.

Bộ nhớ này là một phần khá quan trọng của bộ nhớ trong bởi máy tính có khởi động được hay không là nhờ vào thiết bị này. ROM khác hẳn so với RAM làm cho dữ liệu sẽ không bị mất khi bạn tắt máy, có thể đọc mà không thể thay đổi và sửa chữa.

Một chip ROM có thể lưu trữ được vài megabyte, khi một chip RAM có thể lên đến hàng chục gigabyte.

ROM có một số loại cơ bản sau:

BIOS ROM (Basic Input-Output System Read Only Memory): Là một chip nhớ đặc biệt chứa chương trình nhập xuất cơ sở của hệ thống (BIOS), được nhà sản xuất tích hợp trên bo mạch chủ, giữ vai trò là cầu nối giữa các thiết bị phần cứng với hệ điều hành.

– ROM mặt nạ: Thông tin được ghi khi sản xuất, rất đắt

– PROM (Programmable ROM): là loại chip được lập trình bằng chương trình đặc biệt, dữ liệu sẽ không bị mất khi tắt máy. Được lập trình một lần và dữ liệu trên chip không thể xóa.

– EPROM (Erasable Programmable ROM): Cần thiết bị chuyên dụng để ghi bằng chương trình. Ghi được nhiều lần, trước khi ghi lại, xóa bằng tia cực tím

– EEPROM (Electrically Erasable Programmable ROM): có thể ghi theo từng byte, thông tin có xóa bằng điện.

– Flash memory (Bộ nhớ cực nhanh): Ghi theo khối, thông tin có thể xóa bằng điện.



*Bộ nhớ ROM trên thực tế*

## ⑤ Bộ nhớ RAM

### Bộ nhớ RAM là gì ? Mục đích sử dụng ?

RAM (viết tắt của Random Access Memory) là một loại bộ nhớ khả biến cho phép truy xuất đọc-ghi ngẫu nhiên đến bất kỳ vị trí nào trong bộ nhớ dựa theo địa chỉ bộ nhớ. Thông tin lưu trên RAM chỉ là tạm thời, chúng sẽ mất đi khi mất nguồn điện cung cấp.

RAM là bộ nhớ chính của máy tính và các hệ thống điều khiển, để lưu trữ các thông tin thay đổi đang sử dụng. Các hệ thống điều khiển còn sử dụng SRAM như làm một thiết bị lưu trữ thứ cấp (secondary storage). Khi cần thiết thì bố trí một pin nhỏ làm nguồn điện phụ để duy trì dữ liệu trong RAM. RAM có một đặc tính là thời gian thực hiện thao tác đọc hoặc ghi đối với mỗi ô nhớ là như nhau, cho dù đang ở bất kỳ vị trí nào trong bộ nhớ. Mỗi ô nhớ của RAM đều có một địa chỉ. Thông thường, mỗi ô nhớ là một byte (8 bit); tuy nhiên hệ thống lại có thể đọc ra hay ghi vào nhiều byte (2, 4, 8 byte) một lúc.

RAM khác biệt với các thiết bị bộ nhớ tuần tự (sequential memory device) chẳng hạn như các băng từ, CD-RW, DVD-RW, ổ đĩa cứng, trong đó bắt buộc phải tìm đến sector và đọc/ghi cả khối dữ liệu ở đó để truy xuất. RAM là thuật ngữ phân biệt tương đối theo ý nghĩa sử dụng, với các chip nhớ truy xuất ngẫu nhiên là EEPROM (read-only memory) cấm hoặc hạn chế chiều ghi, và bộ nhớ flash được phép đọc/ghi.

Máy vi tính sử dụng RAM để lưu trữ mã chương trình và dữ liệu trong suốt quá trình thực thi. Đặc trưng tiêu biểu của RAM là có thể truy cập vào những vị trí khác nhau trong bộ nhớ và hoàn tất trong khoảng thời gian tương tự, ngược lại với một số kỹ thuật khác, đòi hỏi phải có một khoảng thời gian trì hoãn nhất định.

### Lịch sử của RAM

Chip RAM có mặt trên thị trường vào cuối những năm 1960, với sản phẩm DRAM là Intel 1103 công bố vào tháng 10 năm 1970.

Những chiếc laptop thế hệ đầu được sản xuất vào cuối những năm thập niên 90 được tích hợp bên trong chúng là RAM SDR với tốc độ khá chậm và bộ nhớ rất ít. Hiện nay khó có thể còn được gặp loại RAM này nữa. Để cải thiện vấn đề tốc độ cũng như bộ nhớ của SDR, người ta bắt đầu nghiên cứu và sản xuất thế hệ RAM tiếp theo vào đầu những năm 2000 với tên gọi mới là DDR, đây cũng là nền móng cho các loại RAM hiện đại sau này. Tuy vậy tốc độ của RAM thế hệ này vẫn còn là rất chậm. Được sử dụng rộng rãi trên các Laptop từ đầu những năm 2000 đến cuối 2004, DDR SDRAM (Double Data Rate SDRAM), thường được giới chuyên môn gọi tắt là "DDR". DDR SDRAM là cải tiến của bộ nhớ SDR với tốc độ truyền tải gấp đôi SDR nhờ vào việc truyền tải hai lần trong một chu kỳ bộ nhớ.

Vào các năm tiếp theo, với sự xuất hiện của các thế hệ máy tính mới hơn, đi kèm với các hệ điều hành có giao diện thân thiện hơn với người dùng, một trong số đó là Windows XP ra đời vào năm 2003. Với sự xuất hiện của các giao diện "màu mè" và "mượt mà" hơn so với thế hệ trước do đó RAM cũng đòi hỏi nhanh hơn và có bộ nhớ lớn hơn. DDR dường như đã quá sức sau khoảng vài năm phục vụ, vì thế người ta nghĩ đến việc thay thế nó bằng thế hệ tiếp theo DDR2 có tốc độ nhanh hơn và bộ nhớ lớn hơn khá nhiều, đồng thời cũng tiết kiệm năng lượng hơn so với DDR. Hiện tại còn có thể tìm thấy DDR2 với bộ nhớ đến 4GB. RAM DDR2 được sử dụng khá nhiều, chúng xuất hiện trên các dòng Laptop sản xuất từ cuối năm 2003 đến cuối năm 2009.

Vào năm 2007 cùng với sự ra đời của các hệ điều hành mới như Windows Vista, Mac OS X Leopard, người ta bắt đầu sản xuất thế hệ RAM tiếp theo của DDR2 là DDR3 với tốc độ rất nhanh và bộ nhớ lớn đồng thời tiết kiệm năng lượng hơn 30% so với thế hệ DDR2. Đây cũng là loại RAM được sử dụng rộng rãi, phổ biến trên Laptop hiện nay với bộ nhớ lớn lên đến 16GB/thanh. Tuy xuất hiện sớm nhưng mãi đến cuối năm 2009 thì DDR3 mới bắt đầu xuất hiện rộng rãi trên Laptop. Bên cạnh sự xuất hiện của DDR3 người ta còn thấy DDR3L. Đây là kết quả hợp tác của Kingston và Intel trong việc phát triển dòng bộ nhớ tiết kiệm điện năng. Chữ L ở đây có nghĩa là Low ám chỉ đây là loại RAM DDR3 nhưng sử dụng ít năng lượng hơn. Loại RAM này thường được sản xuất cho các thiết bị cụ thể vì chúng sử dụng điện thế 1,35V thay vì 1,5V như các loại RAM thông thường. Đây là loại RAM đặc biệt được thiết kế cho các hệ thống máy chủ, các trung tâm dữ liệu và trên một số dòng Laptop cao cấp nhằm tăng thời gian sử dụng pin.

Khi RAM DDR3 vừa tròn 8 năm tuổi đời thì DDR4 cũng được ra mắt. Những thay đổi đáng chú ý nhất của DDR4 so với người tiền nhiệm DDR3 gồm: gia tăng số tuỳ chọn xung nhịp (clock) và chu kỳ (timing), giảm điện năng tiêu thụ (power saving) và giảm độ trễ (latency). Hiện tại, DDR3 đang được giới hạn chủ yếu ở 4 mức xung nhịp 1333, 1600 và 1866 MHz. Mức 2133 MHz đang là mức giới hạn xung nhịp về lý thuyết cho DDR3, trong khi các mức 800 MHz và 1066 MHz giờ đã không còn được tiếp tục sản xuất.

Tháng 11 năm 2018, SK Hynix đã trình làng module DDR5 16GB đầu tiên trên thế giới, hứa hẹn về sự xuất hiện của loại DRAM đang trong quá trình nghiên cứu này.

### Đặc trưng của RAM

Bộ nhớ RAM có 4 đặc trưng sau:

* Dung lượng bộ nhớ: Tổng số byte của bộ nhớ (nếu tính theo byte) hoặc là tổng số bit trong bộ nhớ nếu tính theo bit.
* Tổ chức bộ nhớ: Số ô nhớ và số bit cho mỗi ô nhớ
* Thời gian thâm nhập: Thời gian từ lúc đưa ra địa chỉ của ô nhớ đến lúc đọc được nội dung của ô nhớ đó.
* Chu kỳ bộ nhớ: Thời gian giữa hai lần liên tiếp thâm nhập bộ nhớ.

### Phân loại RAM

Tùy theo công nghệ chế tạo của các nước trên thế giới, được chia làm hai loại.

RAM tĩnh

RAM tĩnh - SRAM (Static Random Access Memory) được chế tạo theo công nghệ ECL (dùng trong CMOS và BiCMOS). Mỗi bit nhớ gồm có các cổng logic với 6 transistor MOS. SRAM là bộ nhớ nhanh, việc đọc không làm hủy nội dung của ô nhớ và thời gian thâm nhập bằng chu kỳ của bộ nhớ. Nhưng SRAM là một nơi lưu trữ các tập tin của CMOS dùng cho việc khởi động máy.

RAM động

RAM động - DRAM (Dynamic Random Access Memory) dùng kỹ thuật MOS. Mỗi bit nhớ gồm một transistor và một tụ điện. Việc ghi nhớ dữ liệu dựa vào việc duy trì điện tích nạp vào tụ điện và như vậy việc đọc một bit nhớ làm nội dung bit này bị hủy. Do vậy sau mỗi lần đọc một ô nhớ, bộ phận điều khiển bộ nhớ phải viết lại nội dung ô nhớ đó. Chu kỳ bộ nhớ cũng theo đó mà ít nhất là gấp đôi thời gian thâm nhập ô nhớ.

Việc lưu giữ thông tin trong bit nhớ chỉ là tạm thời vì tụ điện sẽ phóng hết điện tích đã nạp và như vậy phải làm tươi bộ nhớ sau khoảng thời gian 2μs. Việc làm tươi được thực hiện với tất cả các ô nhớ trong bộ nhớ. Công việc này được thực hiện tự động bởi một vi mạch bộ nhớ.

Bộ nhớ DRAM chậm nhưng rẻ tiền hơn SRAM.

**Các loại DRAM:**

SDRAM

(Viết tắt từ Synchronous Dynamic RAM) được gọi là DRAM đồng bộ. SDRAM gồm 3 phân loại: SDR, DDR, DDR2,DDR3 và DDR4

SDR SDRAM

(Single Data Rate SDRAM), thường được giới chuyên môn gọi tắt là "SDR". Có 168 chân. Được dùng trong các máy vi tính cũ, bus speed chạy cùng vận tốc với clock speed của memory chip, nay đã lỗi thời.

DDR SDRAM

(Double Data Rate SDRAM), thường được giới chuyên môn gọi tắt là "DDR". Có 184 chân. DDR SDRAM là cải tiến của bộ nhớ SDR với tốc độ truyền tải gấp đôi SDR nhờ vào việc truyền tải hai lần trong một chu kỳ bộ nhớ. Đã được thay thế bởi DDR2.

DDR2 SDRAM

(Double Data Rate 2 SDRAM), Thường được giới chuyên môn gọi tắt là "DDR2". Là thế hệ thứ hai của DDR với 240 chân, lợi thế lớn nhất của nó so với DDR là có bus speed cao gấp đôi clock speed.

DDR3 SDRAM

(Double Data Rate III SDRAM): có tốc độ bus 800/1066/1333/1600 Mhz, số bit dữ liệu là 64, điện thế là 1.5v, tổng số pin là 240.

RDRAM

(Viết tắt từ Rambus Dynamic RAM), thường được giới chuyên môn gọi tắt là "Rambus". Đây là một loại DRAM được thiết kế kỹ thuật hoàn toàn mới so với kỹ thuật SDRAM. RDRAM hoạt động đồng bộ theo một hệ thống lặp và truyền dữ liệu theo một hướng. Một kênh bộ nhớ RDRAM có thể hỗ trợ đến 32 chip DRAM. Mỗi chip được ghép nối tuần tự trên một module gọi là RIMM (Rambus Inline Memory Module) nhưng việc truyền dữ liệu được thực hiện giữa các mạch điều khiển và từng chip riêng biệt chứ không truyền giữa các chip với nhau. Bus bộ nhớ RDRAM là đường dẫn liên tục đi qua các chip và module trên bus, mỗi module có các chân vào và ra trên các đầu đối diện. Do đó, nếu các khe cắm không chứa RIMM sẽ phải gắn một module liên tục để đảm bảo đường truyền được nối liền. Tốc độ Rambus đạt từ 400-800 MHz. Rambus tuy không nhanh hơn SDRAM là bao nhưng lại đắt hơn rất nhiều nên có rất ít người dùng. RDRAM phải cắm thành cặp và ở những khe trống phải cắm những thanh RAM giả (còn gọi là C-RIMM) cho đủ.

LPDDR

(Low Power Double Data Rate SDRAM), là loại DRAM có điện năng thấp. Được đóng gói dưới dạng BGA (chân bi), loại DRAM này thường được sử dụng trên các loại điện thoại thông minh, máy tính bảng, laptop siêu mỏng...

### Các thông số của RAM

Được phân loại theo chuẩn của JEDEC.

**Dung lượng**

Dung lượng RAM được tính bằng MB và GB, thông thường RAM được thiết kế với các dung lượng 256mb,512 mb,1gb,2gb,3gb,4gb,8gb... Dung lượng của RAM càng lớn càng tốt cho hệ thống, tuy nhiên không phải tất cả các hệ thống phần cứng và hệ điều hành đều hỗ trợ các loại RAM có dung lượng lớn, một số hệ thống phần cứng của máy tính cá nhân chỉ hỗ trợ đến tối đa 4 GB và một số hệ điều hành (như phiên bản 32 bit của Windows XP) chỉ hỗ trợ đến 32 GB.

**BUS**

Phân loại

Có hai loại BUS là: BUS Speed và BUS Width.

- BUS Speed chính là BUS RAM, là tốc độ dữ liệu được xử lý trong một giây.

- BUS Width là chiều rộng của bộ nhớ. Các loại RAM DDR, DDR2, DDR3, DDR4 hiện nay đều có BUS Width cố định là 64.

Công thức tính băng thông (bandwidth) từ BUS Speed và BUS Width:

*Bandwidth = (Bus Speed x Bus Width) / 8*

- Bandwidth là tốc độ tối đa RAM có thể đọc được trong một giây. Bandwidth được ghi trên RAM là con số tối đa theo lý thuyết. Trên thực tế, bandwidth thường thấp hơn và không thể vượt quá được con số theo lý thuyết.

Các loại RAM, BUS RAM và Bandwidth tương ứng

SDR SDRAM được phân loại theo bus speed như sau:

PC-66: 66 MHz bus.

PC-100: 100 MHz bus.

PC-133: 133 MHz bus.

DDR SDRAM được phân loại theo bus speed và bandwidth như sau:

DDR-200: Còn được gọi là PC-1600. 100 MHz bus với 1600 MB/s bandwidth.

DDR-266: Còn được gọi là PC-2100. 133 MHz bus với 2100 MB/s bandwidth.

DDR-333: Còn được gọi là PC-2700. 166 MHz bus với 2667 MB/s bandwidth.

DDR-400: Còn được gọi là PC-3200. 200 MHz bus với 3200 MB/s bandwidth.

DDR2 SDRAM được phân loại theo bus speed và bandwidth như sau:

DDR2-400: Còn được gọi là PC2-3200. 100 MHz clock, 200 MHz bus với 3200 MB/s bandwidth.

DDR2-533: Còn được gọi là PC2-4200. 133 MHz clock, 266 MHz bus với 4267 MB/s bandwidth.

DDR2-667: Còn được gọi là PC2-5300. 166 MHz clock, 333 MHz bus với 5333 MB/s bandwidth.

DDR2-800: Còn được gọi là PC2-6400. 200 MHz clock, 400 MHz bus với 6400 MB/s bandwidth

DDR3 SDRAM được phân loại theo bus speed và bandwidth như sau:

DDR3-1066: Còn được gọi là PC3-8500. 533 MHz clock, 1066 MHz bus với 8528 MB/s bandwidth

DDR3-1333: Còn được gọi là PC3-10600. 667 MHz clock, 1333 MHz bus với 10664 MB/s bandwidth

DDR3-1600: Còn được gọi là PC3-12800. 800 MHz clock, 1600 MHz bus với 12800 MB/s bandwidth

DDR3-2133: Còn được gọi là PC3-17000. 1066 MHz clock, 2133 MHz bus với 17064 MB/s bandwidth

DDR4 SDRAM được phân loại theo bus speed và bandwidth như sau:

DDR4-2133: Tên module PC4-17000. 1067 MHz clock, 2133 MHz bus với 17064 MB/s bandwidth.

DDR4-2400: Tên module PC4-19200. 1200 MHz clock, 2400 MHz bus với 19200 MB/s bandwidth.

DDR4-2666: Tên module PC4-21300. 1333 MHz clock, 2666 MHz bus với 21328 MB/s bandwidth.

DDR4-3200: Tên module PC4-25600. 1600 MHz clock, 3200 MHz bus với 25600 MB/s bandwidth.

### Các loại module của RAM và tính tương thích bo mạch chủ

Trước đây, các loại RAM được các hãng sản xuất thiết kế cắm các chip nhớ trên bo mạch chủ thông qua các đế cắm (có dạng DIP theo hình minh hoạ trên), điều này thường không thuận tiện cho sự nâng cấp hệ thống. Cùng với sự phát triển chung của công nghệ máy tính, các RAM được thiết kế thành các module như SIMM, DIMM (như hình minh hoạ trên) để thuận tiện cho thiết kế và nâng cấp hệ thống máy tính.

SIMM (Single In-line Memory Module)

DIMM (Dual In-line Memory Module)

SO-DIMM: (Small Outline Dual In-line Memory Module): Thường sử dụng trong các máy tính xách tay.

Không phải các RAM khác nhau đều sử dụng được trên tất cả các bo mạch chủ. Mỗi loại bo mạch chủ lại sử dụng với một loại RAM khác nhau tuỳ thuộc vào chipset của bo mạch chủ.Đó là các bo mạch chủ sử dụng CPU Intel (trước đời Core i) bởi vì trong chipset đó có tích hợp điều khiển bộ nhớ(memory controller). Còn đối hệ thống sử dụng CPU AMD thì việc quản lý bộ nhớ Ram phụ thuộc vào chính CPU. Bởi trong CPU AMD tích hợp điều khiển bộ nhớ (trình điều khiển bộ nhớ) trong chính CPU. Đặc biệt sau này trình điều khiển bộ nhớ đã được tích hợp trong hệ thống Core i của Intel.

