I**.Khái niệm bộ nhớ và lịch sử phát triển của bộ nhớ.**

**Bộ nhớ** là nơi chứa các tiến trình và dữ liệu của tiến trình. Đây là tài nguyên quan trọng thứ hai sau CPU trong một hệ thống máy tính. Bộ nhớ là khối ô nhớ được nhóm lại thành các từ hay các byte và được đánh địa chỉ. Địa chỉ được sử dụng khi cần đọc hoặc ghi thông tin vào bộ nhớ. Trong những hệ điều hành đa nhiệm, nhiều tiến trình có thể cùng thực hiện một lúc và được chứa trong bộ nhớ.

*Vậy là từ những năm 1940 với ống chân không và trống từ, hệ thống bộ nhớ đã trải qua các bước tiến lớn với sự xuất hiện của bộ nhớ bán dẫn, bộ nhớ ngoài, và các công nghệ bộ nhớ mới như SSD và 3D NAND. Ngày nay, bộ nhớ không chỉ tập trung vào dung lượng mà còn cả về tốc độ truy cập, độ bền và khả năng tiết kiệm năng lượng, phục vụ tốt hơn cho các ứng dụng trong máy tính cá nhân, smartphone, và trung tâm dữ liệu.*

**II. Cấu trúc và tổ chức hệ thống bộ nhớ**

1. Các bộ nhớ

Phân chia bộ nhớ máy tính thành 2 loại chính :

* Bộ nhớ trong

+ Bộ nhớ đệm Cache.

+ Bộ nhớ chính (Bộ nhớ trung tâm): RAM, DRAM, ROM, PROM,..

* Bộ nhớ ngoài : băng từ, đĩa (đĩa cứng, đĩa mềm, CD,….)
  1. **Bộ nhớ trong**
* Khái niệm :

Bộ nhớ trong là bộ nhớ được tích hợp trực tiếp và hoạt động trong các thiết bị như máy tính, điện thoại di động hoặc máy tính bảng. Bộ nhớ trong bao gồm: Bộ nhớ chính (RAM,ROM,…) và bộ nhớ đệm (Cache).

- Chức năng :

Chức năng chính của bộ nhớ trong là lưu dữ liệu và thông tin tạm thời, giúp nâng cao hiệu quả hoạt động và trải nghiệm của người sử dụng đồng thời cho phép tiếp cận nhanh chóng các thông tin cần thiết mà không cần sự can thiệp từ thiết bị ngoại vi khác.

*1.1.1 Bộ nhớ trung tâm (hay còn được gọi là Bộ nhớ chính)*

**Bộ nhớ chính** là nơi chứa các chương trình đang được thực hiện, bao gồm cả các lệnh của chương trình cũng như dữ liệu. Bộ nhớ chính bao gồm các loại như : RAM, ROM,…

+ Bộ nhớ RAM (Random Access Memory):

Bộ nhớ RAM là bộ nhớ giúp máy tính lưu trữ dữ liệu ngắn hạn mà các ứng dụng và hệ điều hành đang sử dụng, giúp việc truy cập thông tin diễn ra nhanh chóng. Người ta hay gọi bộ nhớ RAM là bộ nhớ nhanh nhất nhưng lại là bộ nhớ không ổn định. Điều này là vì để giữ lại nội dung, bộ nhớ RAM phải được cung cấp nguồn điện liên tục. Khi nguồn điện cho chip bộ nhớ này bị tắt, chip bộ nhớ này sẽ mất tất cả nội dung. Rất dễ đọc và ghi dữ liệu vào bộ nhớ RAM. Dữ liệu sẽ được đọc hoặc ghi vào bộ nhớ RAM bằng tín hiệu điện.

* RAM có hai dạng cơ bản là DRAM (Dynamic RAM) và SRAM (Static RAM).
* DRAM là bộ nhớ động, đây là bô nhớ không ổn định vì nó sẽ bị mất dữ liệu khi không có nguồn cấp. Trong DRAM, các ô nhớ được tạo bằng tụ điện. Khi tụ điện được tích điện, giá trị của ô nhớ được coi là 1, còn ngược lại là 0. Điều đó có nghĩa là khi tụ điện đã tích điện thì tương ứng là 1, không tích điện thì tương ứng là 0. Tuy nhiên tụ điện sẽ tự động phát điện sau một thời gian nên để giữ lại dữ liệu trong tụ thì phải sạc định kỳ.
* SRAM là bộ nhớ tĩnh, khác với SRAM là nó không cần làm mới liên tục để duy trì dữ liệu. SRAM lưu trữ dữ liệu bằng cách sử dụng các mạch flip-flop, giúp nó tiêu tốn ít năng lượng hơn để duy trì trạng thái, nhưng nếu mất nguồn điện hoàn toàn, dữ liệu trong SRAM cũng sẽ bị mất đi.

Chúng ta cần lưu ý rằng: Cả hai dạng RAM này đều dễ bay hơi vì vậy hãy kết nối nguồn điện liên tục để duy trì giá trị bit của chúng.

+ Bộ nhớ ROM (Read-Only Memory):

Trong các loại bộ nhớ trong của máy tính thì bộ nhớ chỉ đọc (ROM) là bộ nhớ ổn định nhất vì các ô nhớ của chip bộ nhớ này khôn cần nguồn điện để giữ lại giá trị bit của nó. Vì đây là bộ nhớ chỉ đọc nên các giá trị bit của bộ nhớ này chỉ có thể được đọc và không thể ghi hoặc sửa đổi.

Ưu điêm của bộ nhớ trong này là dữ liệu hoặc chương trình cần thiết luôn có trong bộ nhớ chính bên trong và không bắt buộc phải tải dữ liệu từ bất kỳ bộ nhớ phụ nào như trong RAM.

* Giống như RAM, ROM cũng có 1 số dạng khác là PROM, EPROM, EEPROM và bộ nhớ Flash.
* PROM (Chip bộ nhớ chỉ đọc lập trình được) được sử dụng khi cần ít ROM với nội dung bộ nhớ cụ thể. PROM chỉ có thể được viế một lần bằng tín hiệu điện.
* EPROM (Bộ nhớ chỉ đọc có thể lập trình có thể xóa) có thể được đọc và ghi bằng tín hiệu điện. Trước khi thao tác ghi được thực hiện, nội dung bộ nhớ của chip bộ nhớ này sẽ bị xóa để lấy lại trạng thái ban đầu. EPROM có thể bị xóa và cập nhật liên tục và giống như ROM và PROM. Nó vẫn giữ lại nội dung bộ nhớ ngay cả khi không có nguồn điện.
* EEPROM (Chip nhớ không bay hơi) là bộ nhớ ROM có thể được xóa có chọn lọc và ghi nhiều lần. Không giống như EPROM, toàn bộ dữ liệu sẽ bị xóa khi tiếp xúc với tia cực tím. Trong EEPROM, dữ liệu chọn lọc có thể bị xóa mà không cần tháo chip bộ nhớ khỏi hệ thống vì dữ liệu sẽ bị xóa khi có điện áp cao hơn bình thường.

Bộ nhớ Flash là loại EEPROM đặc biệt. Nó có thể ghi toàn bộ khối ô. Trước khi ghi vào bộ nhớ flash, khối ô phải được xóa. So với EEPROM, việc xóa bộ nhớ flash nhanh hơn.

*1.1.2 Bộ nhớ đệm (Cache)*

- Bộ nhớ đệm là bộ nhớ trung gian giữa bộ nhớ chính và CPU, dùng để lưu trữ thông tin tạm thời giúp tăng tốc độ xử lý yêu cầu từ các ứng dụng. Trong nhiều thiết bị như máy tính hay điện thoại, Cache thường được tích hợp trong CPU và có nhiều cấp độ khác nhau như L1, L2, L3 và đôi khi là L4, mỗi cấp có tốc độ và dung lượng khác nhau.

*Tóm lại có một vấn đề là, RAM và bộ nhớ đệm chỉ có thể lưu giữ dữ liệu đến khi mà nó có điện. Với các dữ liệu vẫn còn trên thiết bị khi tắt máy, chúng phải được chuyển vào trong các bộ lưu trữ dài hạn . Đó chính là nhiệm vụ của Bộ nhớ ngoài.*

* 1. *Bộ nhớ ngoài*
* **Khái niệm :**

**Bộ nhớ ngoài** bao gồm các thiết bị lưu trữ dữ liệu không nằm trực tiếp trên bo mạch chủ của máy tính. Khác với bộ nhớ chính (RAM), bộ nhớ ngoài có thể được kết nối với máy tính thông qua các cổng kết nối như USB, HDMI hoặc các kết nối không dây. Bộ nhớ ngoài cung cấp khả năng lưu trữ linh hoạt, dễ dàng sao lưu dữ liệu và mở rộng dung lượng lưu trữ của hệ thống máy tính.

* Phân loại :

Bộ nhớ ngoài thường được chia thành 3 loại :

+ Đĩa từ : lưu trữ bằng từ trường. Dữ liệu được lưu trong một mô hình từ tính trên một đĩa quay được phủ phim từ tính. Nhưng bởi vì phải quay đến nơi dữ liệu được lưu trữ để có thể đọc được, nên độ trễ của các loại ổ này chậm hơn gấp 100000 lần so với DRAM.

+ Đĩa quang : là các bộ lưu trữ trên cơ sở quang học như đĩa DVD cũng dử dụng các đĩa quay nhưng với lớp phủ phản xạ. Chúng được mã hóa bằng các đường dẫn sáng tối trên mặt đĩa và sử dụng các điốt có thể đọc được bằng một tia laser.

+ Flash Disk : là loại ổ lưu trữ lâu dài mới nhất và nhanh nhất, ổ lưu trữ dạng rắn. Chúng không có các bộ phận chuyển động, thay vào đó chúng sử dụng các bóng bán dẫn cổng nối để lưu trữ các bit bằng cách bắt lấy hoặc loại bỏ các hạt mang điện bên trong một cấu trúc được thiết kế đặc biệt.

**2. Tổ chức hệ thống bộ nhớ**

Ngoài bộ nhớ chính (RAM), máy tính còn nhiều dạng bộ nhớ khác như bộ nhớ thanh ghi, bộ nhớ trên đĩa… Lý do phải sử dụng nhiều dạng bộ nhớ là do không có dạng thiết bị nhớ nào thỏa mãn đồng thời các yêu cầu đặt ra về lưu trữ thông tin.  
Bộ nhớ máy tính lý tưởng là bộ nhớ thỏa mãn đồng thời các yêu cầu chính sau:

1) dung lượng lớn;

2) tốc độ truy cập nhanh;

3) giá thành thấp;

4) có khả năng lưu trữ lâu bền cả khi có điệnvà không có điện.

Những yêu cầu này là mâu thuẫn với nhau, chẳng hạn thiết bị nhớ tốc độ  
cao có giá cao và không lưu được thông tin khi không có điện. Cụ thể, bộ nhớ chính, được xây dựng dựa trên công nghệ DRAM (dynamic random-access memory), mặc dù có tốc độ truy cập tương đối cao xong không đủ lớn để lưu trữ thường xuyên tất cả chương trình và dữ liệu. Ngoài ra, nội dung bộ nhớ sử dụng DRAM sẽ bị xóa khi tắt nguồn, do vậy không phù hợp để lưu trữ lâu thông tin lâu dài.  
 Để giải quyết vấn đề nói trên, hệ thống bộ nhớ trong máy tính được tạo thành từ nhiều dạng bộ nhớ khác nhau, mỗi dạng có ưu điểm về một mặt nào đó như tốc độ, dung lượng, giá thành. Hệ thống bộ nhớ được tổ chức như một cấu trúc phân cấp hình tháp như minh họa trên hình 2.1, trong đó các dạng bộ nhớ ở mức trên có tốc độ và giá thành cao, do vậy chỉ có thể sử dụng với dung lượng nhỏ. Ngược lại, phía dưới của tháp là bộ nhớ dung lượng lớn và rẻ nhưng chậm.

Các dạng bộ nhớ từ SSD trở xuống có thể lưu trữ thông tin và dữ liệu ngay cả khi không có điện, trong khi các dạng bộ nhớ phía trên trong sơ đồ phân cấp bị mất nội dung khi không có nguồn nuôi. Bộ nhớ SSD là dạng bộ nhớ sử dụng công nghệ nhớ mới, có tốc độ nhanh hơn đĩa từ, trong khi vẫn có thể lưu thông tin khi không có điện. Dạng bộ nhớ SSD thông dụng nhất là dạng bộ nhớ flash dùng cho các thiết bị như máy ảnh, điện thoại di động thông minh.  
 Một số máy tính xách tay (laptop) mới cũng sử dụng bộ nhớ loại này kết hợp với đĩa cứng, trong đó bộ nhớ SSD được sử dụng cho để lưu những thông tin cần truy cập nhanh như thông tin dùng để chuyển máy từ trạng thái “ngủ” sang trạng thái hoạt động. Do giá thành đang giảm đi nhanh trong khi dung lượng ngày càng lớn nên bộ nhớ SSD được sử dụng ngày càng phổ biến.

**3. So sánh cấu trúc tổ chức hoạt động hệ thống bộ nhớ của máy tính cá nhân PC với cấu trúc tổ chức hệ thống của bộ nhớ của Smartphone**

**3.1 Giống nhau**

*3.1.1 Cấu trúc phân cấp:*

Cả hai hệ thống đều sử dụng cấu trúc bộ nhớ phân cấp, bao gồm:

**+ Bộ nhớ đệm (Cache):** Nằm gần CPU nhất, có tốc độ truy cập nhanh nhất, lưu trữ dữ liệu được sử dụng thường xuyên.

**+ Bộ nhớ chính (RAM):** Dung lượng lớn hơn Cache, tốc độ truy cập chậm hơn, lưu trữ dữ liệu và chương trình đang được sử dụng.

**+ Bộ nhớ trong:** Dung lượng lớn nhất, tốc độ truy cập chậm nhất, lưu trữ dữ liệu và chương trình lâu dài.

**3.1.2. Bộ nhớ ảo:**

Cả PC và smartphone đều sử dụng bộ nhớ ảo để mở rộng không gian địa chỉ bộ nhớ vật lý. Khi RAM đầy, hệ điều hành sẽ chuyển một phần dữ liệu ít sử dụng từ RAM sang bộ nhớ trong (ổ cứng hoặc bộ nhớ flash), tạo ra không gian trống cho các ứng dụng mới.

**3.1.3. Quản lý bộ nhớ:**

Hệ điều hành trên cả hai thiết bị đều có nhiệm vụ quản lý bộ nhớ, bao gồm:

**+ Phân bổ bộ nhớ:** Cấp phát bộ nhớ cho các chương trình và dữ liệu.

**+ Giải phóng bộ nhớ:** Thu hồi bộ nhớ khi không còn sử dụng.

**+ Bảo vệ bộ nhớ:** Ngăn chặn các chương trình truy cập vào vùng nhớ không được phép.

**3.1.4. Mục đích sử dụng:**

Về cơ bản, mục đích sử dụng bộ nhớ trên cả hai thiết bị là tương tự nhau:

**+ Lưu trữ hệ điều hành:** Chứa hệ điều hành và các tệp tin hệ thống cần thiết để khởi động và vận hành thiết bị.

**+ Chạy ứng dụng:** Cung cấp không gian để tải và chạy các ứng dụng.

**+ Lưu trữ dữ liệu:** Lưu trữ dữ liệu người dùng như hình ảnh, video, tài liệu...

**3.2 Khác nhau:**

**3.2.1. Tương tác CPU và bộ nhớ:**

A white and black text on a white background

Description automatically generated

**3.2.2. Vai trò của bộ nhớ đệm:**

A white text on a white background

Description automatically generated

**3.2.3. Bộ nhớ chính (RAM):**

A white sheet with black text

Description automatically generated

**3.2.4. Truy cập bộ nhớ trong:**

A white text with black text

Description automatically generated

**3.2.5. Quản lý bộ nhớ ảo:**

* **PC và smartphone:**

**+ Kỹ thuật:** Cả hai đều sử dụng kỹ thuật "paging" hoặc "swapping" để chuyển dữ liệu giữa RAM và bộ nhớ trong.

**+ Thuật toán:** Sử dụng các thuật toán thay thế trang như FIFO (First-In, First-Out), LRU (Least Recently Used) để quyết định trang nào sẽ được hoán đổi ra khỏi RAM.

**+ Hiệu năng:** Hiệu năng của bộ nhớ ảo phụ thuộc vào tốc độ của bộ nhớ trong. Do đó, PC với ổ cứng SSD sẽ có hiệu năng bộ nhớ ảo tốt hơn smartphone với bộ nhớ flash.

**3.2.6. Tối ưu hóa cho tác vụ:**

A white text with black text

Description automatically generated

**III. Chương trình mô tả thuật toán lựa chọn vùng trống để cấp phát theo chiến lược first fit, best fit và worst fit gồm 7 vùng nhớ có kích thước khác nhau và kích thước vùng nhớ được cấp phát có giá trị nhập từ bàn phím.**

1. **Giới thiệu chung**

Khi có một tiến trình cần được nạp vào bộ nhớ mà bộ nhớ có nhiều hơn một khối nhớ trống có kích thước lớn hơn kích thước của tiến trình đó, Hệ điều hành phải quyết định chọn một khối nhớ phù hợp để nạp tiến trình sao cho việc lựa chọn này dẫn đến việc sử dụng bộ nhớ chính là hiệu quả nhất.

* Khi sử dụng thuật toán lực chọn vùng trống để cấp phát chúng ta sẽ cấp phát theo các chiến lược first-fit, best-fit, worst-fit.
* **First-fit** : chọn vùng trống đầu tiên có kích thước lớn hơn hoặc bằng kích thước cần cấp phát. Việc tìm kiếm có thể bắt đầu từ đầu danh sách các vùng trống hay từ vị trí của lần cấp phát cuối cùng. Chúng ta dừng tìm kiếm ngay sau khi tìm thấy một vùng trống đủ lớn.  
  **Best-fit** : chọn vùng trống nhỏ nhất trong số các vùng trống có kích thước lớn hơn hoặc bằng kích thước cần cấp phát. Vùng trống sinh ra sau khi cấp phát do vậy sẽ có kích thước bé nhất
* **Worst-fit**: từ nhận xét là các vùng trống sinh ra sau khi cấp phát theo chiến lược thứ hai (best fit) có kích thước bé và do đó thường không thích hợp cho việc cấp phát tiếp theo, người ta nghĩ ra chiến lược thứ ba này. Vùng trống lớn nhất sẽ được cấp phát. Không gian còn thừa từ vùng trống này sau khi cấp xong tạo ra vùng trống mới có kích thước lớn hơn so với hai chiến lược trên.

Ví dụ : Giả sử ta có các vùng nhớ có các vùng nhớ với kích thước :

* Danh sách vùng nhớ : 100KB, 500KB, 200KB, 300KB, 600KB, 400KB, 700KB
* Yêu cầu cấp phát : 350KB (nhập từ bàn phím)

Kết quả :

***Kết quả :***- **First Fit**: Duyệt từ vùng đầu tiên và thấy 500KB là vùng nhớ đầu tiên thỏa mãn.  
-> Dừng lại chương trình ngay  
- **Best-fit**:Sau khi duyệt hết danh sách, vùng nhỏ nhất trong các vùng đủ lớn để cấp phát là 400KB.  
-> Dừng chương trình

- **Worst-fit** : Sau khi duyệt hết danh sách, vùng lớn nhất trong các vùng đủ lớn để cấp phát là 700KB.  
-> Dừng chương trình   
***Kết luận*** : Sau khi thực hiện chương trình, ba chiến lược cấp phát ở trên sẽ cho kết quả như sau

Chiến lược First Fit sẽ chọn vùng nhớ 500KB để chia và cấp phát  
Chiến lược Best Fit sẽ chọn vùng nhớ 400KB để chia và cấp phát

Chiến lược Worst Fit sẽ chọn vùng nhớ 700KB để chia và cấp phát

Để quyết định chiến lược nào là tốt nhất, nhiều nghiên cứu đã được thực hiện. Kết quả mô phỏng và thực nghiệm cho thấy, hai phương pháp đầu cho phép giảm phân mảnh ngoài tốt hơn phương pháp thứ ba. Trong hai phương pháp đầu, phương pháp thứ nhất đơn giản và có tốc độ nhanh nhất

CODE CHƯƠNG TRÌNH : (ĐỌC HIỂU PHẢI LÀM GÌ THÔI)

// Ngôn ngữ sử dụng : Java

import java.util.\*;

public class FitTypes {

public static void main(String[] args) {

Scanner sc = new Scanner(System.in);

ArrayList<Double> a = new ArrayList<Double>();

System.out.print("Nhap kich thuoc vung nho duoc yeu cau cap phat (don vi: MB): ");

double n = sc.nextDouble();

for(int i=0; i<7; ++i){

System.out.print("Nhap kich thuoc vung trong thu " + (i+1) + " cua bo nho: ");

Double b = sc.nextDouble();

a.add(b);

}

double firstFit = -1;

double bestFit = -1;

double worstFit = -1;

// firstFit chương trình duyệt qua a để tìm vùng nhớ đầu tiên có kích thước >=n rồi gán kích thước của vùng đó cho firstFit

for(Double x : a)

if (x>=n){

firstFit = x;

break;

}

// với mỗi vùng trong a nếu kích thước > worstFit và vẫn thỏa mã x>= n , worstFit được cập nhật bằng kích thước của vùng đó

for(Double x : a)

if(x>=n && x>worstFit)

worstFit = x;

Double c = Collections.max(a); // khởi tạo c là phần tử lớn nhất trong a

// chương trình duyệt qua a để tìm bộ nhớ nhỏ nhất có thể chứa n, với mỗi vùng nếu kích thước >=n và<=c thì bestFit được cập nhật

for(Double x : a){

if(x>=n && x<=c){

bestFit = x;

c = x;

}

}

if (firstFit>=n && bestFit>=n && worstFit>=n){

System.out.println("Chien luoc First fit se chon khoi co dung luong " + firstFit + " MB de chia va cap phat");

System.out.println("Chien luoc Best fit se chon khoi co dung luong " + bestFit + " MB de chia va cap phat");

System.out.println("Chien luoc Worst fit se chon khoi co dung luong " + worstFit + " MB de chia va cap phat");

}

else

System.out.println("Khong co vung trong nao thao man yeu cau");

}

}