

Kiến trúc máy tính

Computer Architecture

Giảng viên: Nguyễn Bá Nghiễn, PhD

Bộ môn: Kỹ thuật và mạng máy tính

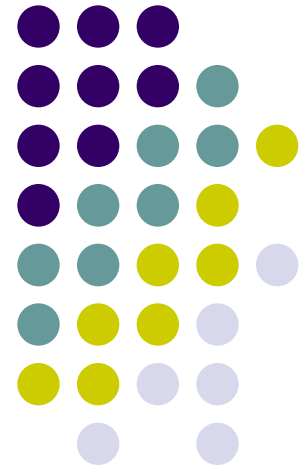
Khoa: Công nghệ thông tin

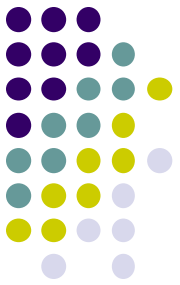
Trường: Đại học công nghiệp hà nội

Điện thoại: 0981687231

Email: nguyen76@gmail.com

nguyenbanghien_cntt@hau.edu.vn

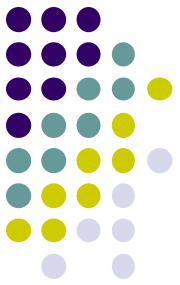




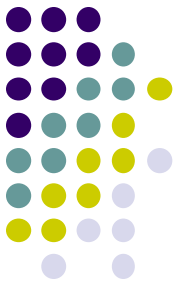
Nội dung môn học

1. Kiến trúc máy tính
2. Kiến trúc bộ lệnh
3. Đường truyền, CPU và hệ thống vào/ra
4. Kiến trúc hệ thống nhớ
5. Kỹ thuật đường ống và RISC

Chương 1. Kiến trúc máy tính



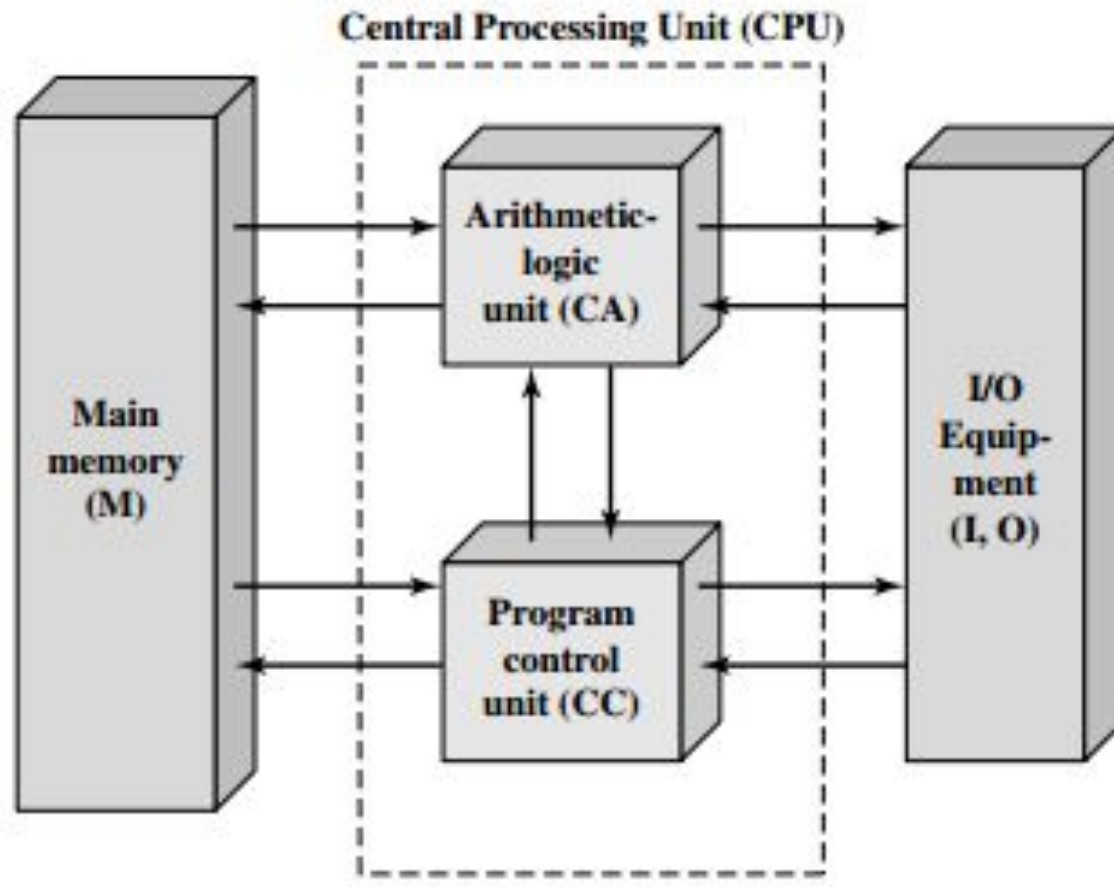
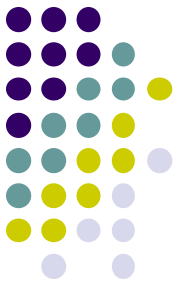
1. Lịch sử phát triển
2. Phân loại kiến trúc máy tính
3. Các dạng máy tính
4. Hệ thống máy tính



Lịch sử phát triển

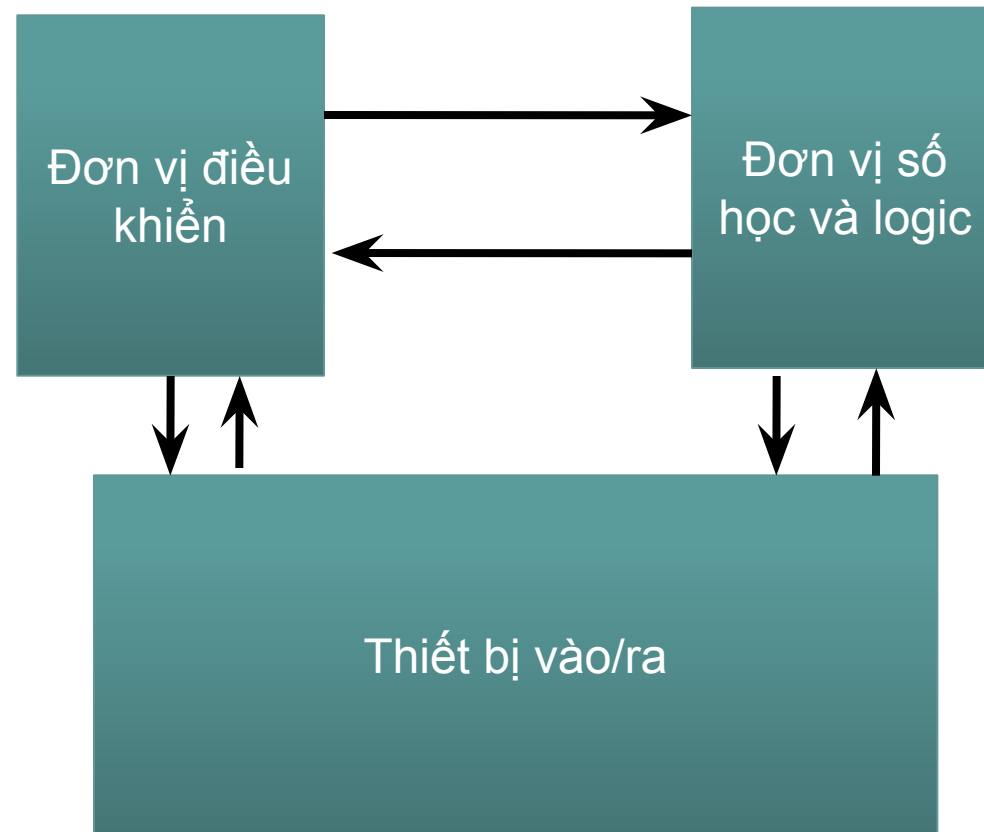
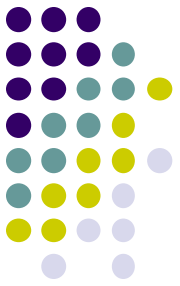
- Thế hệ 1 (1940 – 1950) sử dụng đèn chân không.
- Thế hệ 2 (1950 – 1964) sử dụng Transistor.
- Thế hệ 3 (1964 – 1974) sử dụng mạch tích hợp (IC).
- Thế hệ 4 (1974 – nay) sử dụng mạch tích hợp mật độ cao.

Phân loại kiến trúc máy tính

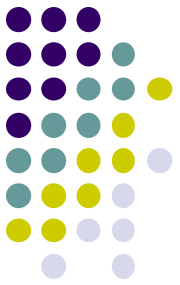


Kiến trúc Von Neumann

Phân loại kiến trúc máy tính (tiếp)



Kiến trúc phi Von Neumann



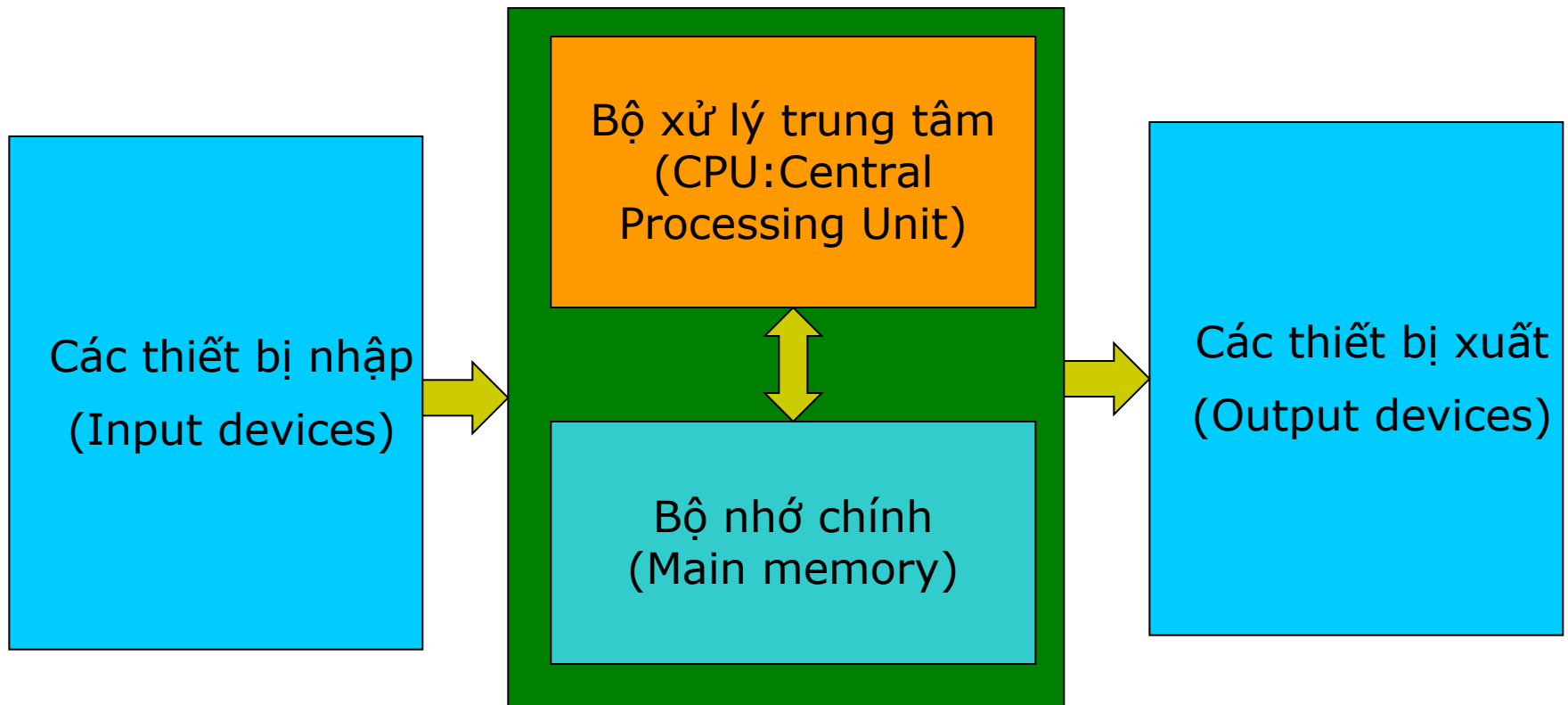
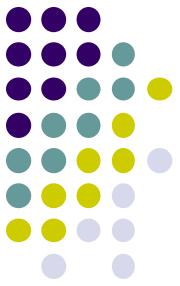
Các dạng máy tính

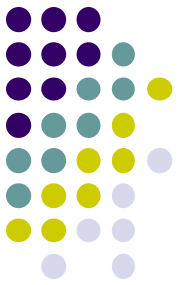
1. Máy tính

- **Máy tính (Computer)** là thiết bị điện tử thực hiện các công việc sau:
 - Nhận thông tin vào,
 - Xử lý thông tin theo dãy các lệnh được nhớ sẵn bên trong,
 - Đưa thông tin ra.
- Dãy các lệnh nằm trong bộ nhớ để yêu cầu máy tính thực hiện công việc cụ thể gọi là **chương trình (program)**

→ Máy tính hoạt động theo chương trình.

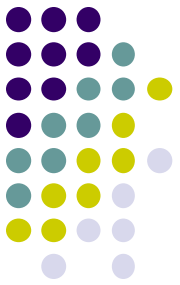
Máy tính (continue)





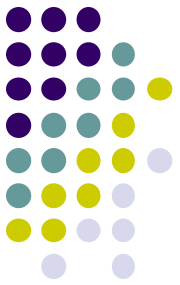
Phân loại máy tính

- Phân loại truyền thống:
 - Máy vi tính (Microcomputers)
 - Máy tính nhỏ (Minicomputers)
 - Máy tính lớn (Mainframe Computers)
 - Siêu máy tính (Supercomputers)



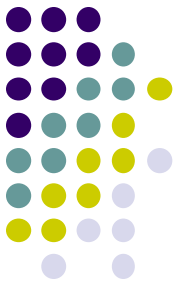
Phân loại máy tính hiện đại

- Máy tính cá nhân (Personal Computers)
- Máy chủ (Server Computers)
- Máy tính nhúng (Embedded Computers)



Máy tính cá nhân PC

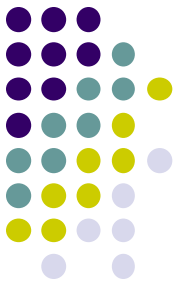
- Là loại máy tính phổ biến nhất
- Các loại máy tính cá nhân:
 - Máy tính để bàn (Desktop)
 - Máy tính xách tay (Laptop)
- 1981 → IBM giới thiệu máy tính IBM-PC sử dụng bộ xử lý Intel 8088
- 1984 → Apple đưa ra Macintosh sử dụng bộ xử lý Motorola 68000
- Giá thành: hàng trăm đến hàng nghìn USD



Máy chủ(server)

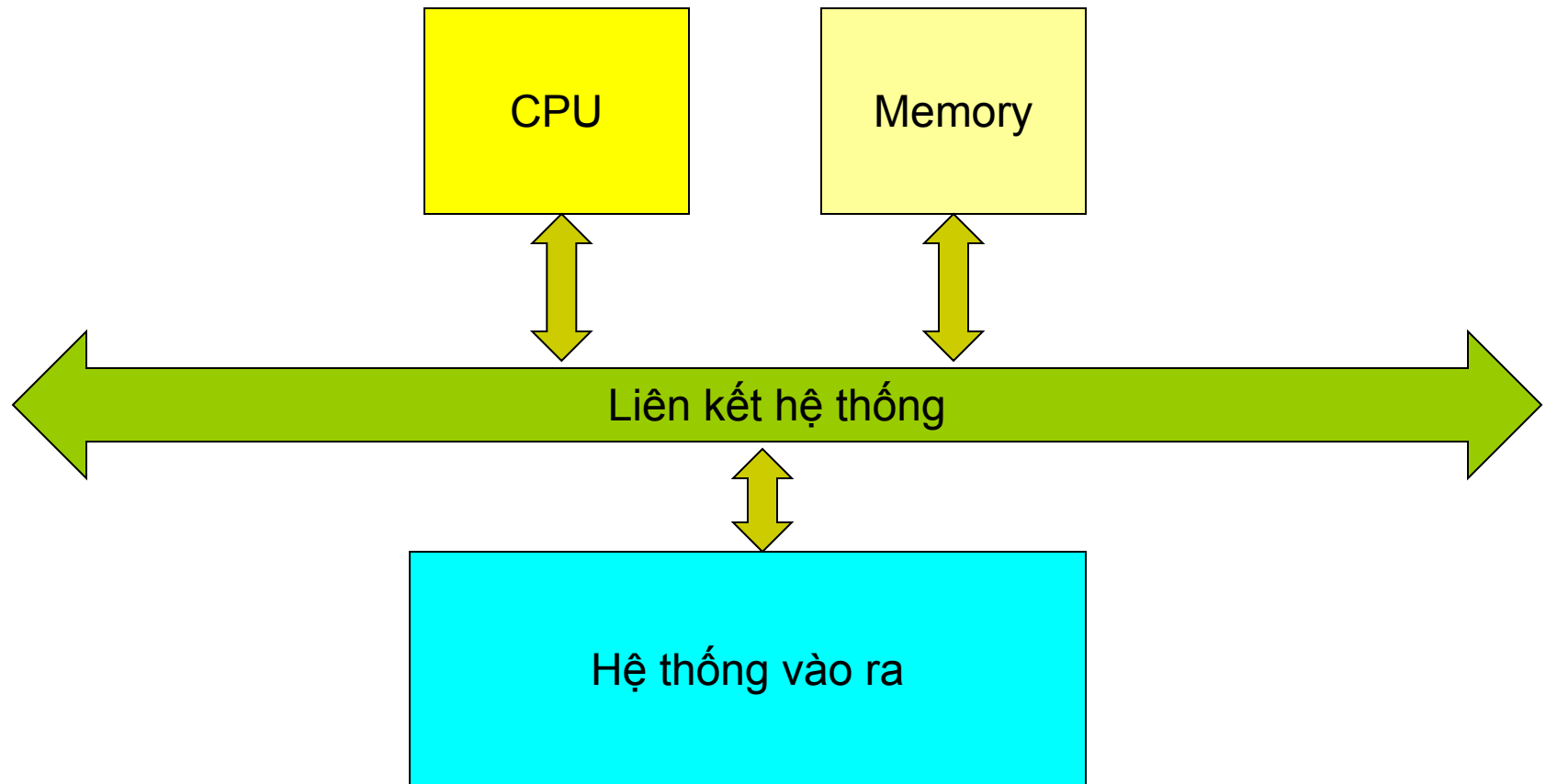
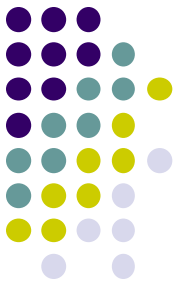
- Thực chất là máy phục vụ
- Dùng trong mạng theo mô hình Client/Server (Khách hàng/Người phục vụ)
- Tốc độ và hiệu năng tính toán cao
- Dung lượng bộ nhớ lớn
- Độ tin cậy cao
- Giá thành: hàng nghìn đến hàng chục triệu USD.

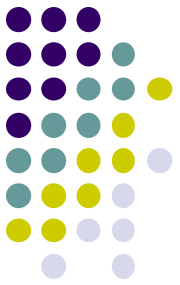
Máy tính nhúng (Embedded Computer)



- Được đặt trong thiết bị khác để điều khiển thiết bị đó làm việc
- Được thiết kế chuyên dụng
- Ví dụ:
 - Điện thoại di động
 - Máy ảnh số
 - Bộ điều khiển trong máy giặt, điều hoà nhiệt độ
 - Router – bộ định tuyến trên mạng
- Giá thành: vài USD đến hàng trăm nghìn USD.

2. Hệ thống máy tính

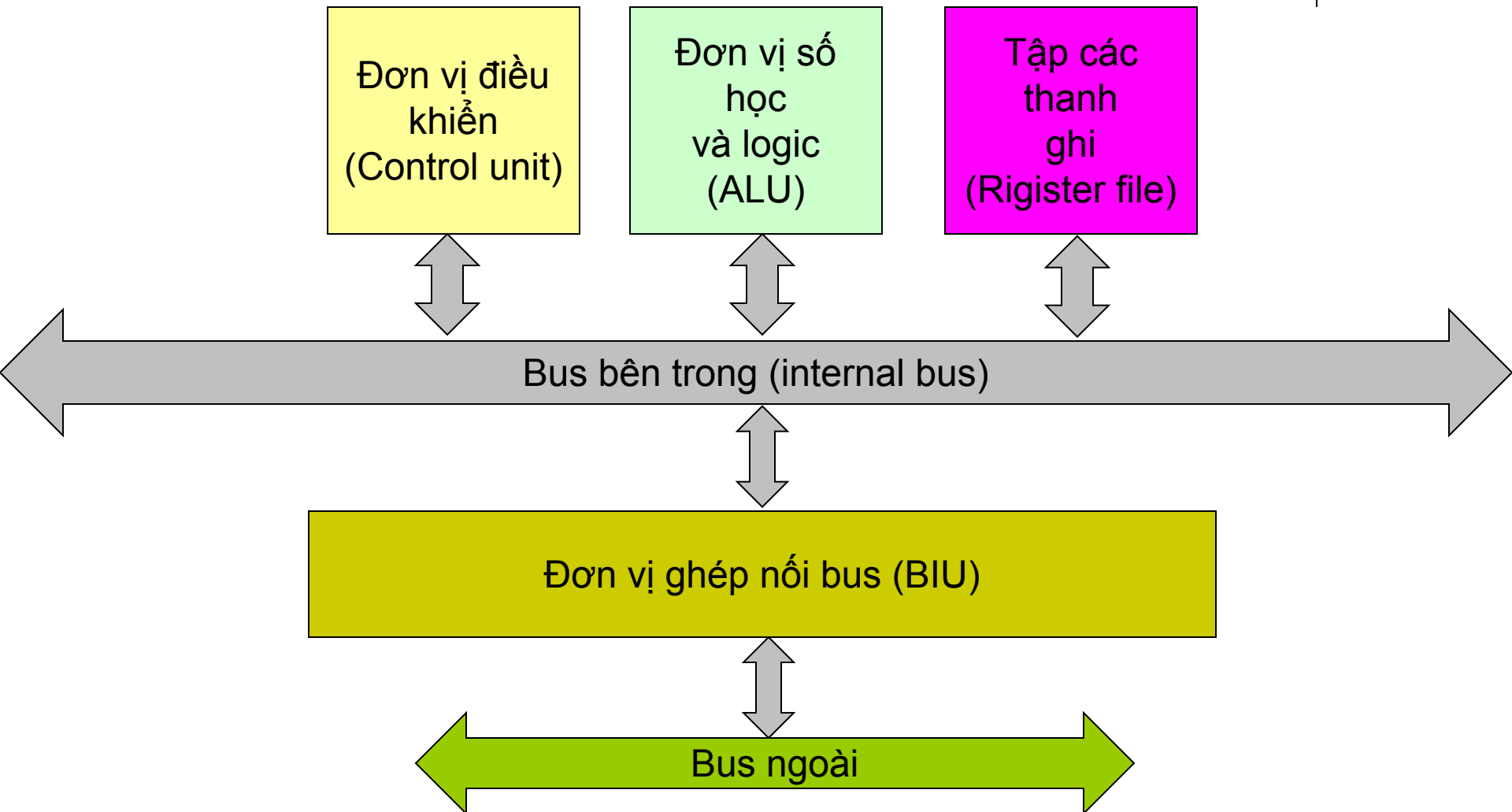
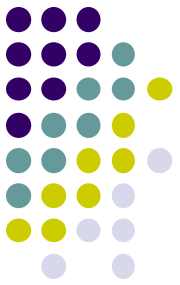




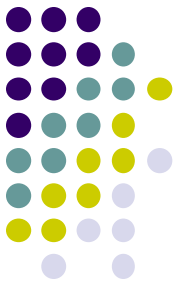
Bộ xử lý trung tâm

- Chức năng:
 - điều khiển hoạt động của máy tính
 - xử lý dữ liệu
- Nguyên tắc hoạt động cơ bản:
CPU hoạt động theo chương trình nằm trong bộ nhớ chính.

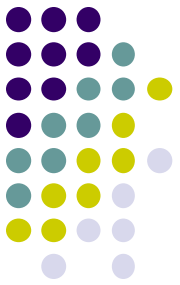
Cấu trúc cơ bản của CPU



Các thành phần cơ bản của CPU

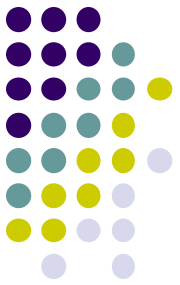


- **Đơn vị điều khiển** (*Control Unit - CU*): điều khiển hoạt động của máy tính theo chương trình đã định sẵn.
- **Đơn vị số học và logic** (*Arithmetic and Logic Unit - ALU*): thực hiện các phép toán số học và phép toán logic.
- **Tập thanh ghi** (*Register File - RF*): lưu giữ các thông tin tạm thời phục vụ cho hoạt động của CPU.
- **Đơn vị nối ghép bus** (*Bus Interface Unit - BIU*) kết nối và trao đổi thông tin giữa bus bên trong (*internal bus*) và bus bên ngoài (*external bus*).

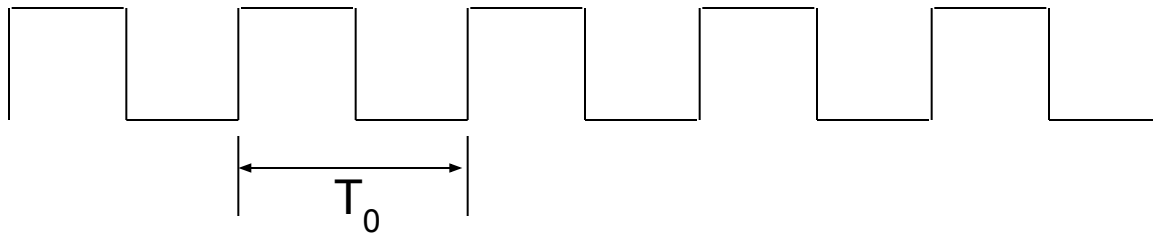


Tốc độ của bộ vi xử lý

- Tốc độ của bộ xử lý:
 - Số lệnh được thực hiện trong 1 giây
 - MIPS (Million of Instructions per Second)
 - Khó đánh giá chính xác
- Tần số xung nhịp của bộ xử lý:
 - Bộ xử lý hoạt động theo một xung nhịp (Clock) có tần số xác định
 - Tốc độ của bộ xử lý được đánh giá gián tiếp thông qua tần số của xung nhịp



Tốc độ của bộ vi xử lý (tiếp)

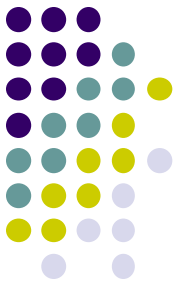


T_0 : chu kỳ xung nhịp

- Tần số xung nhịp: $f_0 = 1/T_0$
- Mỗi thao tác của bộ xử lý cần kT_0
- T_0 càng nhỏ \rightarrow bộ xử lý chạy càng nhanh
- Ví dụ: Máy tính dùng bộ xử lý 2GHz

Ta có $f_0 = 2\text{GHz} = 2 \times 10^9 \text{ (Hz)}$

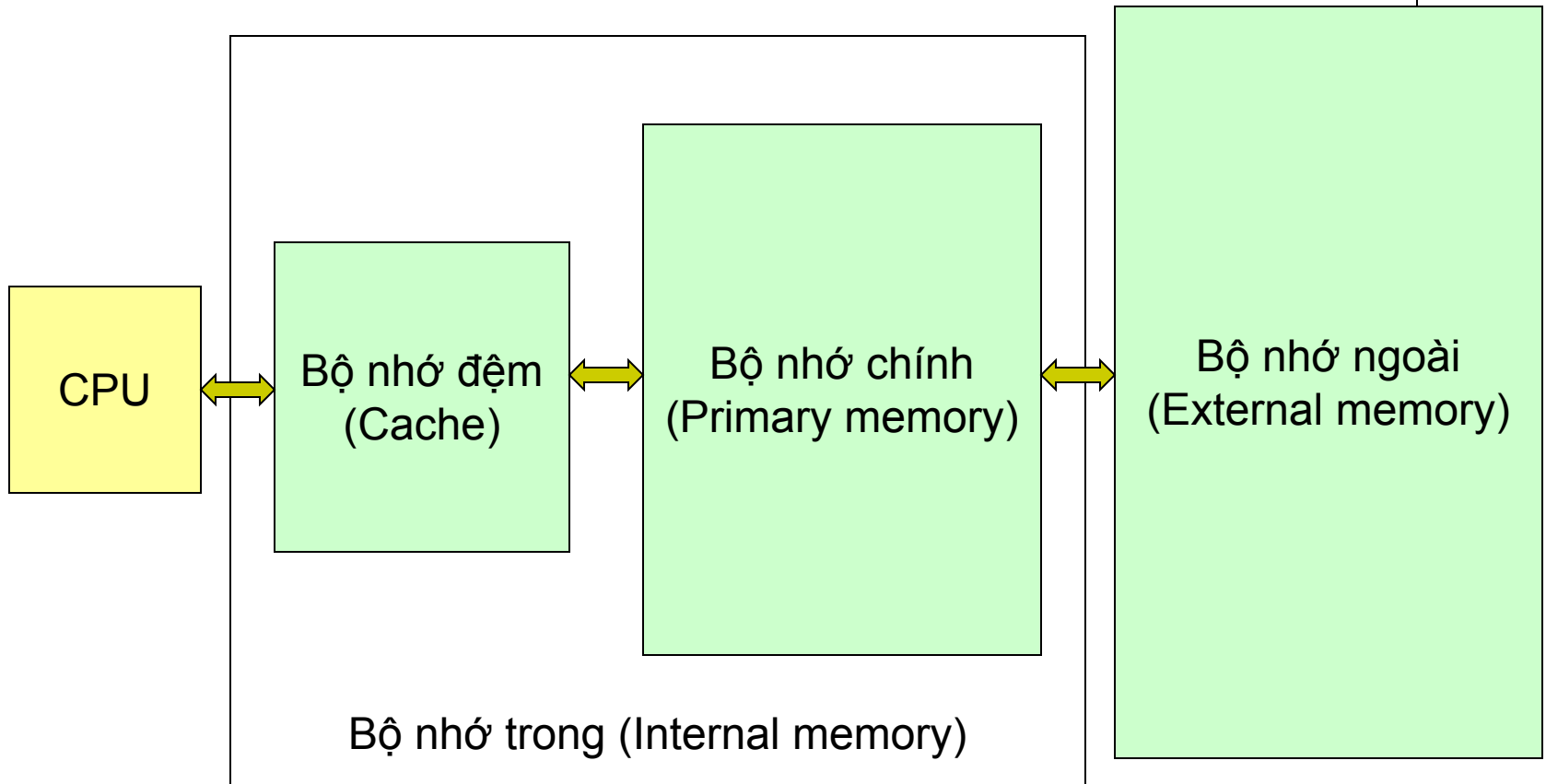
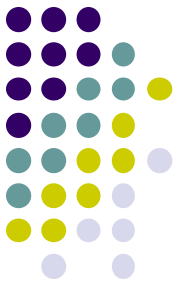
$\rightarrow T_0 = 1/f_0 = 1/(2 \times 10^9) = 0,5 \text{ (ns)}$.

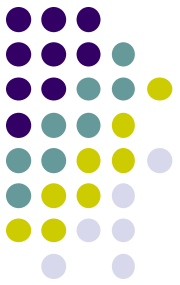


Bộ nhớ của máy tính

- Chức năng: lưu trữ chương trình và dữ liệu.
- Các thao tác cơ bản với bộ nhớ:
 - Thao tác ghi (Write)
 - Thao tác đọc (Read)
- Các thành phần chính:
 - Bộ nhớ trong (Internal Memory)
 - Bộ nhớ ngoài (External Memory)

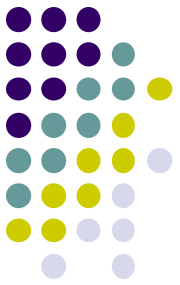
Sơ đồ phân cấp bộ nhớ





Bộ nhớ trong

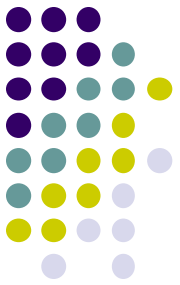
- Chức năng và đặc điểm:
 - Chứa các thông tin mà CPU có thể trao đổi trực tiếp
 - Tốc độ rất nhanh
 - Dung lượng không lớn
 - Sử dụng bộ nhớ bán dẫn: ROM và RAM
- Các loại bộ nhớ trong:
 - Bộ nhớ chính (primary memory)
 - Bộ nhớ đệm (bộ nhớ cache)



Bộ nhớ chính (Primary Memory)

- Chứa các chương trình và dữ liệu đang được CPU sử dụng.
- Tổ chức thành các ngăn nhớ được đánh địa chỉ.
- Ngăn nhớ thường được tổ chức theo byte.
- Nội dung của ngăn nhớ có thể thay đổi, song địa chỉ vật lý của ngăn nhớ luôn cố định.

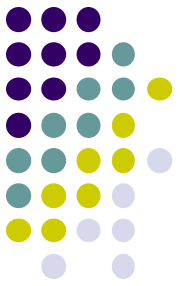
Địa chỉ	Nội dung
0000	10101010
0001	01010000
0002	11110000
0003	00010010
0004	10100110
0005	10111110



Bộ nhớ đệm cache

- Bộ nhớ có tốc độ nhanh được đặt đệm giữa CPU và bộ nhớ chính nhằm tăng tốc độ CPU truy cập bộ nhớ.
- Dung lượng nhỏ hơn bộ nhớ chính.
- Tốc độ nhanh hơn.
- Cache thường được chia thành một số mức.
- Cache có thể được tích hợp trên chip vi xử lý.
- Cache có thể có hoặc không.

Bộ nhớ ngoài (External memory)



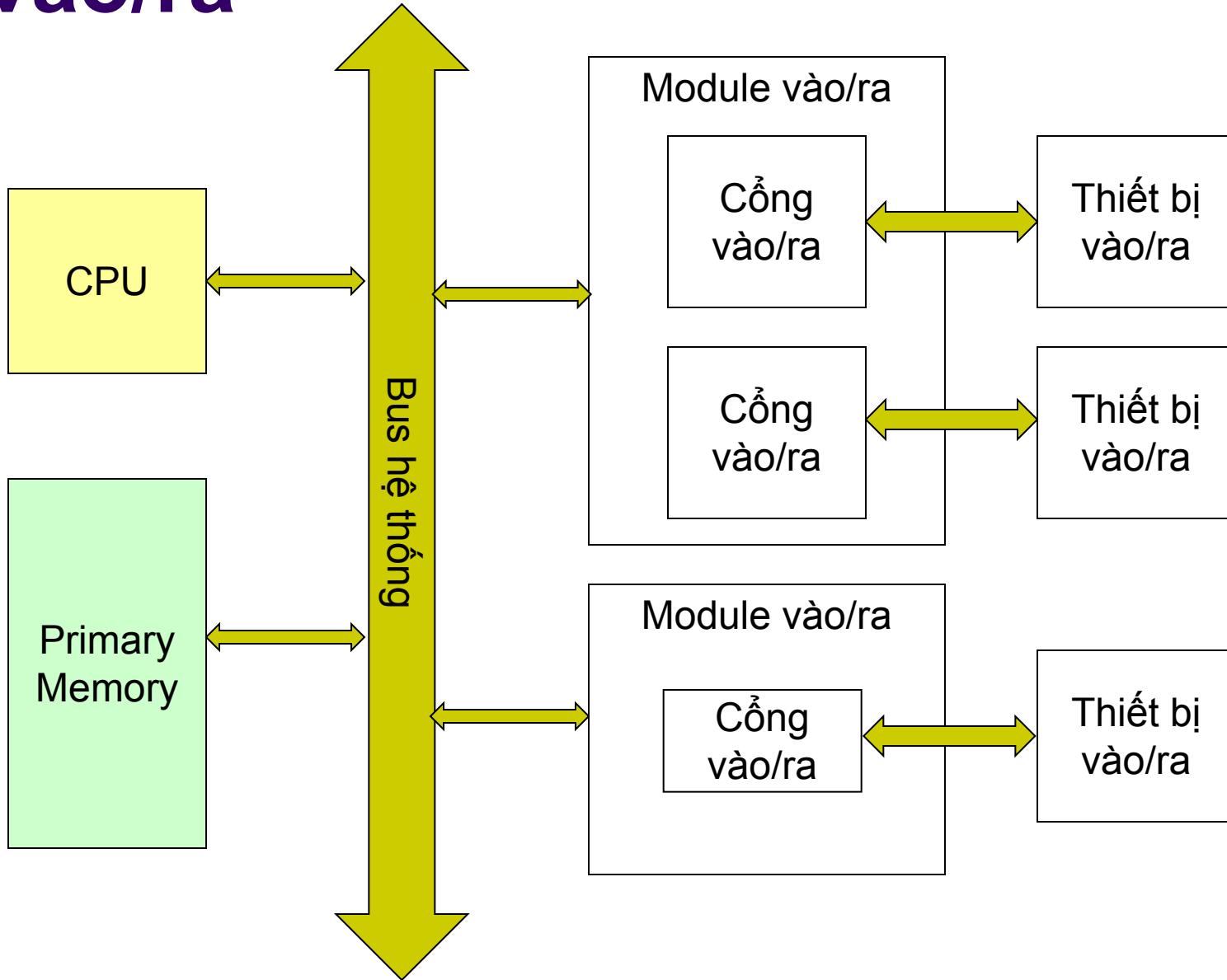
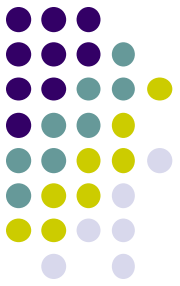
- Chức năng và đặc điểm
 - Lưu giữ tài nguyên phần mềm của máy tính
 - Được kết nối với hệ thống dưới dạng các thiết bị vào-ra
 - Dung lượng lớn
 - Tốc độ chậm
- Các loại bộ nhớ ngoài
 - Bộ nhớ từ: đĩa cứng, đĩa mềm
 - Bộ nhớ quang: đĩa CD, DVD
 - Bộ nhớ bán dẫn: Flash disk, memory card

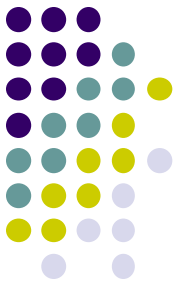
Hệ thống vào/ra (Input/output system)



- Chức năng: Trao đổi thông tin giữa máy tính với thế giới bên ngoài.
- Các thao tác cơ bản:
 - Vào dữ liệu (Input)
 - Ra dữ liệu (Output)
- Các thành phần chính:
 - Các thiết bị ngoại vi (Peripheral Devices)
 - Các mô-đun vào-ra (IO Modules)

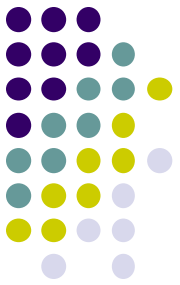
Cấu trúc cơ bản của hệ thống vào/ra





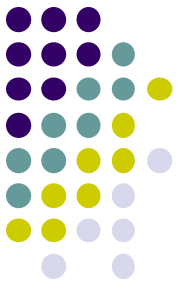
Các thiết bị ngoại vi

- Chức năng: chuyển đổi dữ liệu giữa bên trong và bên ngoài máy tính
- Các loại thiết bị ngoại vi cơ bản
 - Thiết bị vào: bàn phím, chuột, máy quét ...
 - Thiết bị ra: màn hình, máy in ...
 - Thiết bị nhớ: các ổ đĩa ...
 - Thiết bị truyền thông: MODEM ...



Module vào/ra

- Chức năng: nối ghép các thiết bị ngoại vi với máy tính
- Mỗi module vào/ra có một hoặc một vài cổng vào/ra (I/O Port).
- Mỗi cổng vào/ra được đánh một địa chỉ xác định.
- Các thiết bị ngoại vi được kết nối và trao đổi dữ liệu với máy tính thông qua các cổng vào/ra.

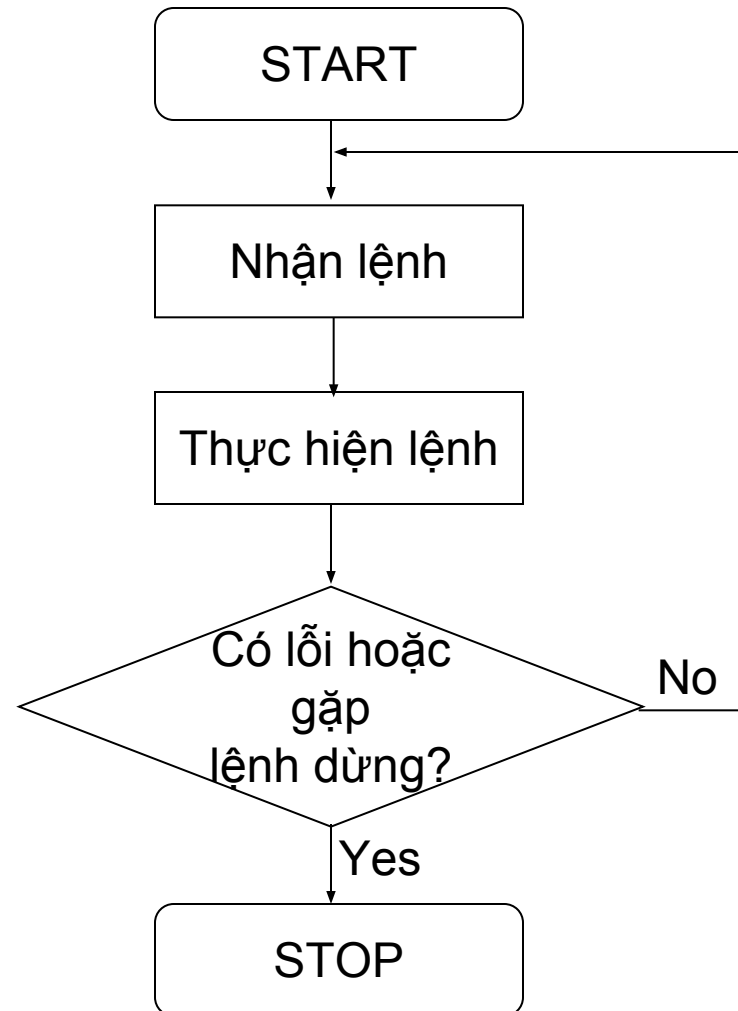
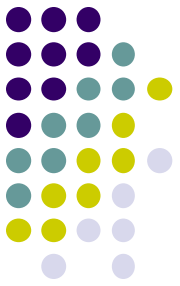


Hoạt động của máy tính

- Thực hiện lệnh: Là hoạt động cơ bản của máy tính
- Máy tính lặp đi lặp lại hai bước:
 - Nhận lệnh
 - Thực hiện lệnh

Chu trình lệnh
- Thực hiện chương trình bị dừng nếu thực hiện lệnh bị lỗi hoặc gặp lệnh dừng.

Chu trình lệnh

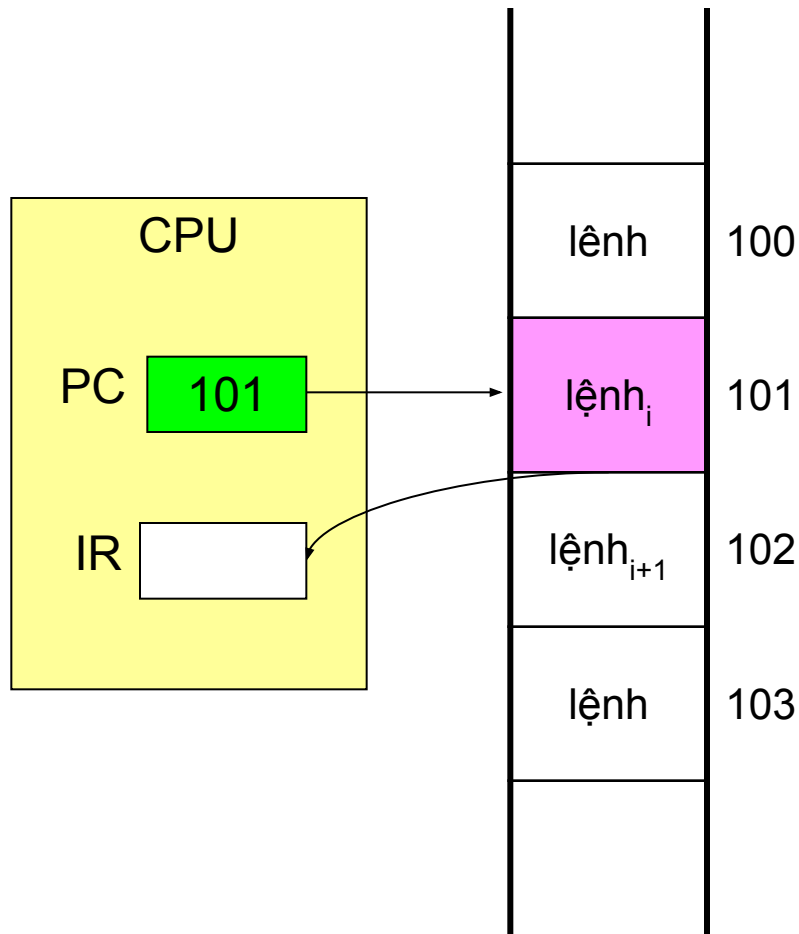
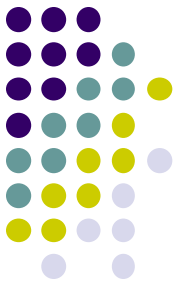




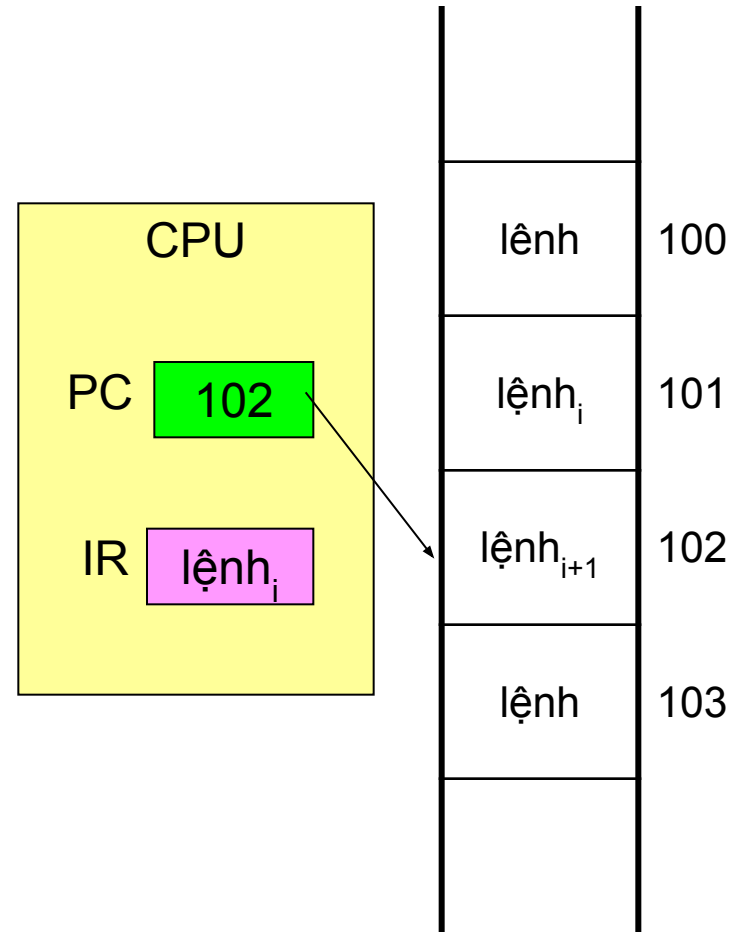
Nhận lệnh

- Bắt đầu mỗi chu trình lệnh, CPU nhận lệnh từ bộ nhớ chính.
- Bộ đếm chương trình PC (Program Counter) của CPU giữ địa chỉ của lệnh sẽ được nhận.
- CPU nhận lệnh từ ngăn nhớ được trả bởi PC.
- Lệnh được nạp vào thanh ghi lệnh IR (Instruction Register).
- Sau khi lệnh được nhận vào, nội dung PC tự động tăng để trả sang lệnh kế tiếp.

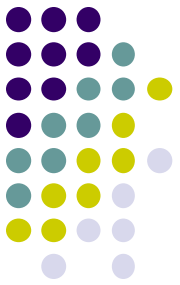
Minh họa quá trình nhận lệnh



Trước khi nhận lệnh_i

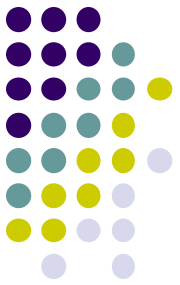


Sau khi nhận lệnh_i



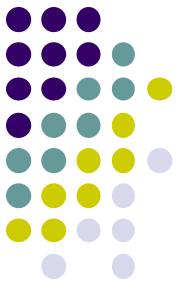
Thực hiện lệnh

- Bộ xử lý giải mã lệnh đã được nhận và phát tín hiệu điều khiển thực hiện thao tác mà lệnh yêu cầu.
- Các kiểu thao tác của lệnh:
 - Trao đổi dữ liệu giữa CPU và bộ nhớ chính
 - Trao đổi dữ liệu giữa CPU và mô-đun vào-ra
 - Xử lý dữ liệu: thực hiện các phép toán số học hoặc phép toán logic với các dữ liệu.
 - Điều khiển rẽ nhánh
 - Kết hợp các thao tác trên.



Ngắt (Interrupt)

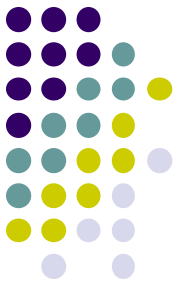
- Khái niệm chung về ngắt: Ngắt là cơ chế cho phép CPU tạm dừng chương trình đang thực hiện để chuyển sang thực hiện một chương trình khác, gọi là *chương trình con phục vụ ngắt (ISR: Interrupt Service Routine)*.
- Các loại ngắt:
 - Ngắt do lỗi khi thực hiện chương trình, ví dụ: tràn số, chia cho 0.
 - Ngắt do lỗi phần cứng, ví dụ lỗi bộ nhớ RAM.
 - Ngắt do mô-đun vào-ra phát tín hiệu ngắt đến CPU yêu cầu trao đổi dữ liệu.



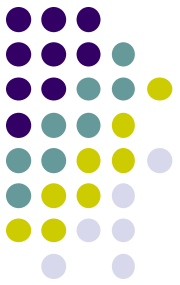
Hoạt động ngắt

- Sau khi hoàn thành mỗi một lệnh, bộ xử lý kiểm tra tín hiệu ngắt
- Nếu không có ngắt thì bộ xử lý nhận lệnh tiếp theo của chương trình hiện tại
- Nếu có tín hiệu ngắt:
 - Tạm dừng chương trình đang thực hiện
 - Cắt ngữ cảnh (các thông tin liên quan đến chương trình bị ngắt)
 - Thiết lập PC trỏ đến chương trình con phục vụ ngắt
 - Chuyển sang thực hiện chương trình con phục vụ ngắt
 - Cuối chương trình con phục vụ ngắt, khôi phục ngữ cảnh và tiếp tục chương trình đang bị tạm dừng

Xử lý nhiều tín hiệu ngắt đồng thời

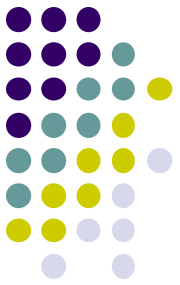


- Xử lý ngắt tuần tự
 - Khi một ngắt đang được thực hiện, các ngắt khác sẽ bị cấm.
 - Bộ xử lý sẽ bỏ qua các ngắt tiếp theo trong khi đang xử lý một ngắt
 - Các yêu cầu ngắt vẫn đang đợi và được kiểm tra sau khi ngắt đầu tiên được xử lý xong
 - Các ngắt được thực hiện tuần tự
- Xử lý ngắt ưu tiên
 - Các ngắt được định nghĩa mức ưu tiên khác nhau
 - Ngắt có mức ưu tiên thấp hơn có thể bị ngắt bởi ngắt ưu tiên cao hơn
 - Xảy ra ngắt lồng nhau



Hoạt động vào/ra

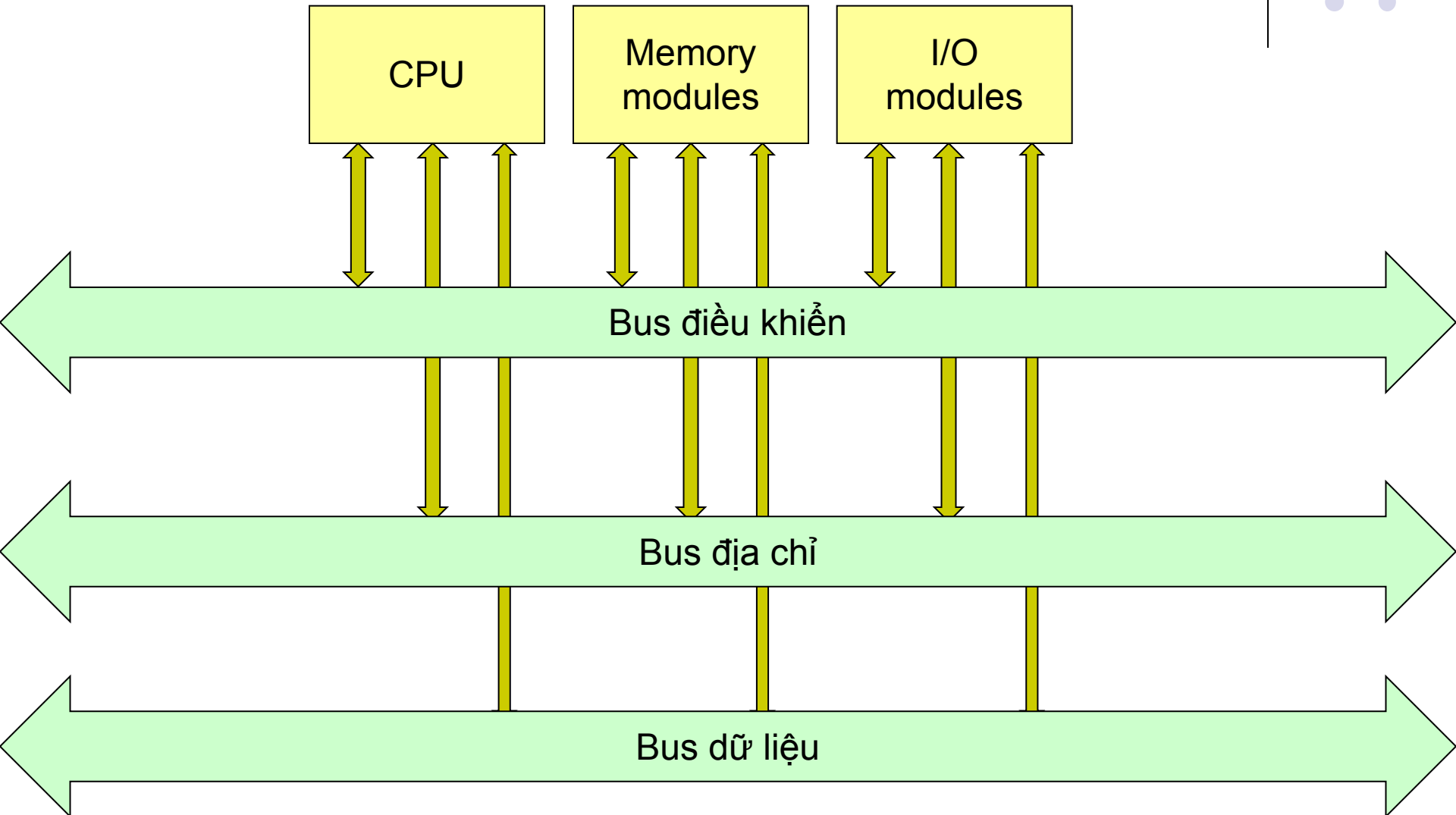
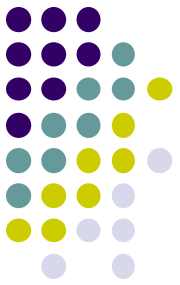
- Hoạt động vào/ra: là hoạt động trao đổi dữ liệu giữa module vào/ra với bên trong máy tính.
- Các kiểu hoạt động vào-ra:
 - CPU trao đổi dữ liệu với module vào/ra
 - Module vào/ra trao đổi dữ liệu trực tiếp với bộ nhớ chính (DMA- Direct Memory Access).



Cấu trúc bus cơ bản

- Bus: tập hợp các đường kết nối dùng để vận chuyển thông tin giữa các module của máy tính với nhau.
- Các bus chức năng:
 - Bus địa chỉ
 - Bus dữ liệu
 - Bus điều khiển
- Độ rộng bus: là số đường dây của bus có thể truyền các bit thông tin đồng thời (chỉ dùng cho bus địa chỉ và bus dữ liệu)

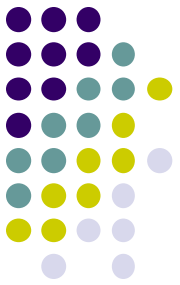
Sơ đồ cấu trúc bus cơ bản





Bus địa chỉ (Address bus)

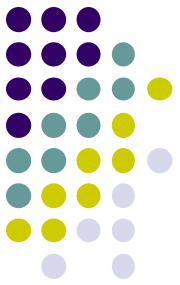
- **Chức năng:** vận chuyển địa chỉ để xác định ngăn nhớ hay cổng vào/ra
- **Độ rộng bus địa chỉ:** cho biết số lượng ngăn nhớ tối đa được đánh địa chỉ.
 - N bit: $A_{N-1}, A_{N-2}, \dots, A_2, A_1, A_0 \rightarrow$ có thể đánh địa chỉ tối đa cho 2^N ngăn nhớ (không gian địa chỉ bộ nhớ)
- **Ví dụ:**
 - Bộ xử lý Pentium có bus địa chỉ 32 bit có khả năng đánh địa chỉ cho 2^{32} bytes nhớ (4GBytes) (ngăn nhớ tổ chức theo byte)



Bus dữ liệu (Data bus)

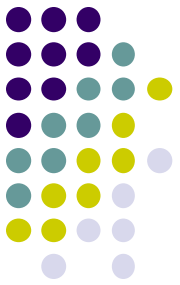
- **Chức năng:**
 - vận chuyển lệnh từ bộ nhớ đến CPU
 - vận chuyển dữ liệu giữa CPU, module nhớ, module vào/ra với nhau
- **Độ rộng bus dữ liệu:** Xác định số bit dữ liệu có thể được trao đổi đồng thời.
 - M bit: $D_{M-1}, D_{M-2}, \dots, D_2, D_1, D_0$
 - M thường là 8, 16, 32, 64, 128 bit.
- **Ví dụ:** Các bộ xử lý Pentium có bus dữ liệu 64 bit

Bus điều khiển (Control bus)



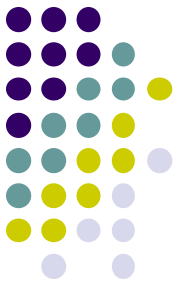
- **Chức năng:** vận chuyển các tín hiệu điều khiển
- **Các loại tín hiệu điều khiển bao gồm:**
 - Các tín hiệu điều khiển đọc/ghi
 - Các tín hiệu điều khiển ngắt
 - Các tín hiệu điều khiển bus

Một số tín hiệu điều khiển thông dụng



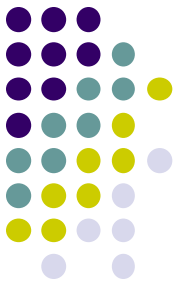
- Các tín hiệu (phát ra từ CPU) điều khiển đọc-ghi:
 - *Memory Read* (MEMR): điều khiển đọc dữ liệu từ một ngăn nhớ có địa chỉ xác định lên bus dữ liệu.
 - *Memory Write* (MEMW): điều khiển ghi dữ liệu có sẵn trên bus dữ liệu đến một ngăn nhớ có địa chỉ xác định.
 - *I/O Read* (IOR): điều khiển đọc dữ liệu từ một cổng vào-ra có địa chỉ xác định lên bus dữ liệu.
 - *I/O Write* (IOW): điều khiển ghi dữ liệu có sẵn trên bus dữ liệu ra một cổng có địa chỉ xác định.

Một số tín hiệu điều khiển thông dụng (tiếp)



- Các tín hiệu điều khiển ngắt:
 - *Interrupt Request* (INTR): Tín hiệu từ bộ điều khiển vào-ra gửi đến yêu cầu ngắt CPU để trao đổi vào/ra. Tín hiệu INTR có thể bị che.
 - *Interrupt Acknowledge* (INTA): Tín hiệu phát ra từ CPU báo cho bộ điều khiển vào-ra biết CPU chấp nhận ngắt để trao đổi vào/ra.
 - *Non Maskable Interrupt* (NMI): tín hiệu ngắt không che được gửi đến ngắt CPU.
 - *Reset*: Tín hiệu từ bên ngoài gửi đến CPU và các thành phần khác để khởi động lại máy tính.

Một số tín hiệu điều khiển thông dụng (tiếp)



- Các tín hiệu điều khiển bus:
 - *Bus Request* (BRQ) hay là *Hold*: Tín hiệu từ module điều khiển vào/ra gửi đến yêu cầu CPU chuyển nhượng quyền sử dụng bus.
 - *Bus Grant* (BGT) hay là *Hold Acknowledge* (HLDA): Tín hiệu phát ra từ CPU chấp nhận chuyển nhượng quyền sử dụng bus.
 - *Lock/ Unlock*: Tín hiệu *cấm/cho phép* xin chuyển nhượng bus



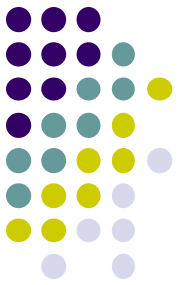
Đặc điểm của cấu trúc đơn bus

- Bus hệ thống chỉ phục vụ được một yêu cầu trao đổi dữ liệu tại một thời điểm
- Bus hệ thống phải có tốc độ bằng tốc độ bus của module chậm nhất trong hệ thống
- Bus hệ thống phụ thuộc vào cấu trúc bus (các tín hiệu) của bộ xử lý → các module nhớ và các module vào/ra cũng phụ thuộc vào bộ xử lý.
- Khắc phục: phân cấp bus → cấu trúc đa bus



Phân cấp bus

- Tổ chức thành nhiều bus trong hệ thống máy tính
 - Cho các thành phần khác nhau:
 - Bus của bộ xử lý
 - Bus của bộ nhớ chính
 - Các bus vào/ra
 - Các bus khác nhau về tốc độ
- Bus bộ nhớ chính và các bus vào/ra không phụ thuộc vào bộ xử lý cụ thể.



Một số bus thông dụng

- Bus của bộ xử lý (Front Side Bus - FSB): có tốc độ nhanh nhất.
- Bus của bộ nhớ chính (Back Side Bus - BSB) (nối ghép với các module RAM).
- AGP bus (Accelerated Graphic Port) - Bus tăng tốc đồ họa: nối ghép card màn hình tăng tốc.
- PCI bus (Peripheral Component Interconnect): nối ghép với các thiết bị ngoại vi có tốc độ trao đổi dữ liệu nhanh.
- IDE (Integrated Device Electronics): Bus kết nối với ổ đĩa cứng hoặc ổ đĩa CD, DVD.
- USB (Universal Serial Bus): Bus nối tiếp đa năng.