

Kiến trúc máy tính

Nhập môn

NGUYỄN Ngọc Hoá

Bộ môn Hệ thống thông tin, Khoa CNTT
Trường Đại học Công nghệ,
Đại học Quốc gia Hà Nội

Nhập môn

- Khái niệm chung
 - Kiến trúc và tổ chức
 - Cấu trúc và chức năng
- Kiến trúc Von Neuman
- Phân loại kiến trúc
- Lịch sử phát triển
- Hiệu năng và luật Moore
- Một số kiến trúc máy tính hiện đại
 - Kiến trúc x86 và x86-64
 - Kiến trúc UltraSparc
- Tổng kết và bài tập

Khái niệm chung

- Kiến trúc máy tính
 - Kiến trúc : hướng đến mục tiêu xây dựng sản phẩm từ những thành phần đã có theo một phương thức nào đó
 - Ví dụ: kiến trúc chip, kiến trúc mạng Internet, kiến trúc hệ thống email, ...
- Wikipedia:
 - “computer architecture is the practical art of selecting and interconnecting hardware components to create computers that meet functional, performance and cost goals and the formal modelling of those systems”.
 - “The coordination of abstract levels of a processor under changing forces, involving design, measurement and evaluation. It also includes the overall fundamental working principle of the internal logical structure of a computer system”...

Architecture & Organization

- Architecture: thiết kế logic của máy tính, đặc tả bởi những thuộc tính mà người lập trình có thể sử dụng được
 - Tập lệnh, cơ chế vào/ra, số bits dùng để biểu diễn dữ liệu...
 - VD: Có lệnh Nhân hay không?
- Organization: thiết kế vật lý của máy tính, đặc tả bởi những cách thức cài đặt các chức năng, thao tác
 - Tín hiệu điều khiển, giao diện, công nghệ nhớ.
 - VD: Bộ nhân đã được cài đặt trực tiếp hay thông qua một dãy các phép cộng?

Mathematics
Algorithms
Applications
Operating Systems
Architecture
Organization
Digital Logic
VLSI Design
Semiconductor Manf.
Physics

Kiến trúc và tổ chức ...

- Họ Intel x86 chia sẻ cùng một kiến trúc cơ bản



- Họ Ultrasparc cũng chia sẻ cùng một kiến trúc cơ bản



- Tuy nhiên, tổ chức của chúng lại khác nhau giữa các phiên bản

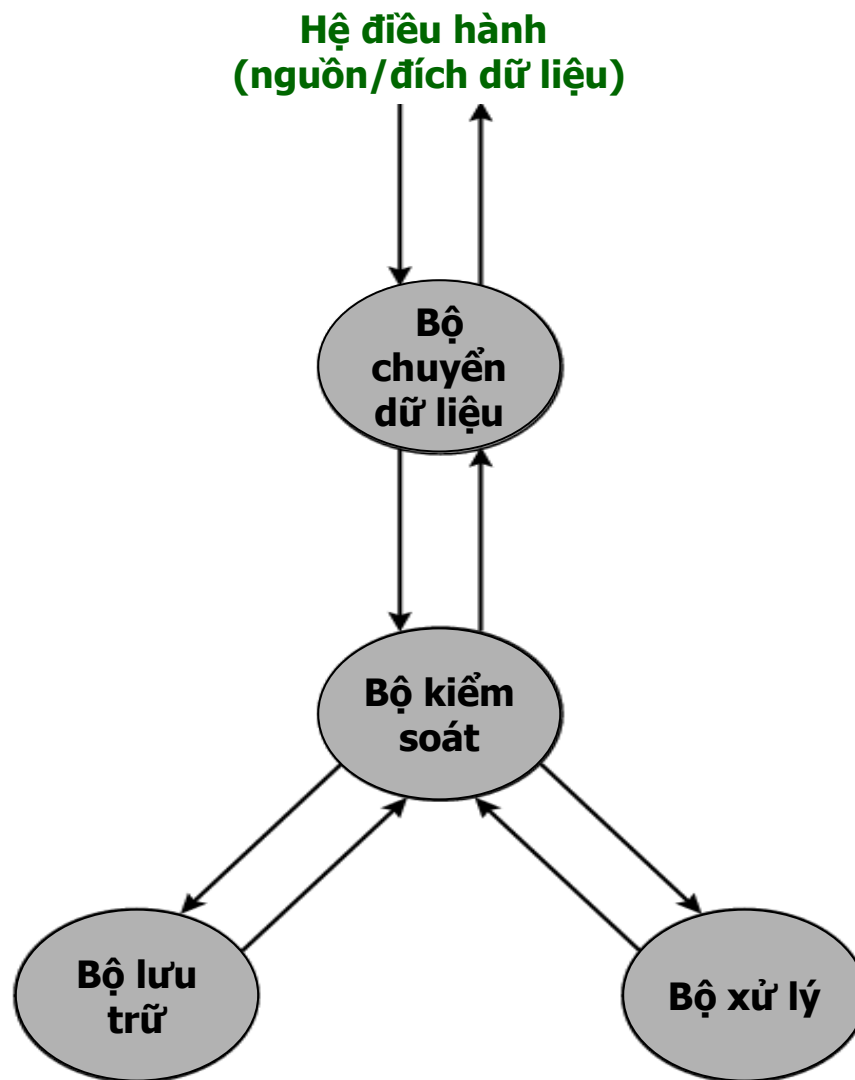
Vai trò của kiến trúc máy tính

- Computer Architecture cho phép:
 - Khai thác những thế mạnh công nghệ
 - Sản xuất các thiết bị nhanh hơn, bé hơn, rẻ hơn, xạch hơn (tiêu thụ ít năng lượng)....
 - Cải thiện được hiệu năng của hệ thống (latency, throughput).
 - Mang đến những ứng dụng mới, khả năng mới
- Sự phát triển của kiến trúc máy tính được xem là mấu chốt cho sự phát triển của những lĩnh vực tính toán khác!

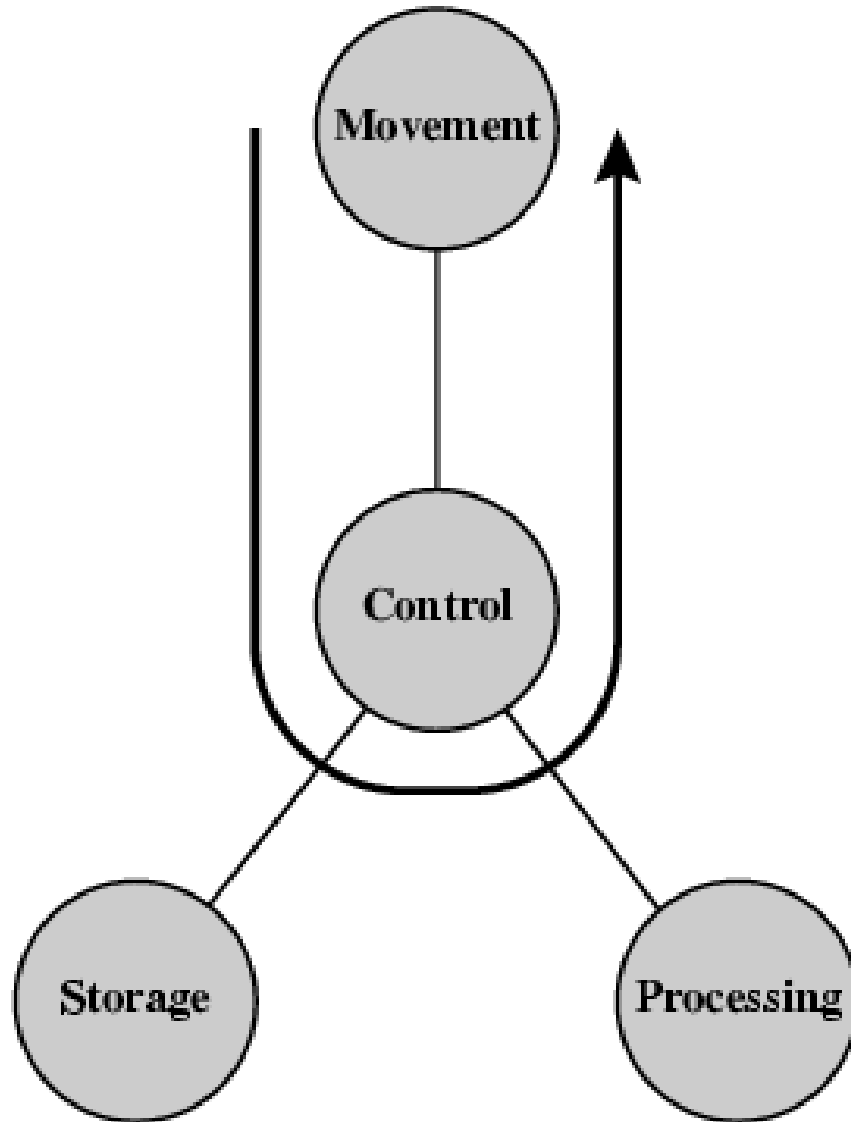
Structure & Function

- Cấu trúc: cách để các thành phần liên kết với nhau
- Chức năng: hoạt động của các thành phần cá nhân trong một phần của cấu trúc
- Chức năng của mọi máy tính:
 - Data processing - Xử lý dữ liệu
 - Data Storage - Lưu giữ dữ liệu
 - Data Movement - Chuyển dữ liệu
 - Control - Điều khiển

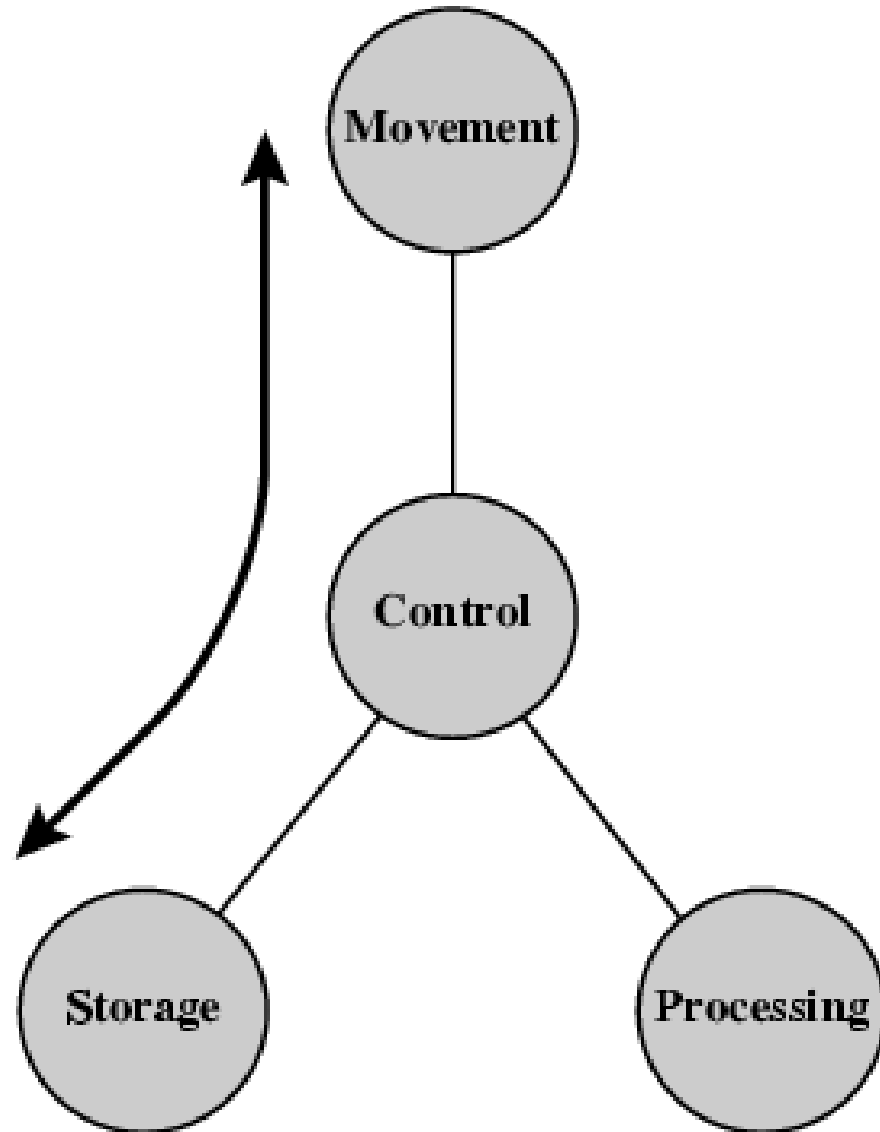
Mối liên hệ giữa các chức năng



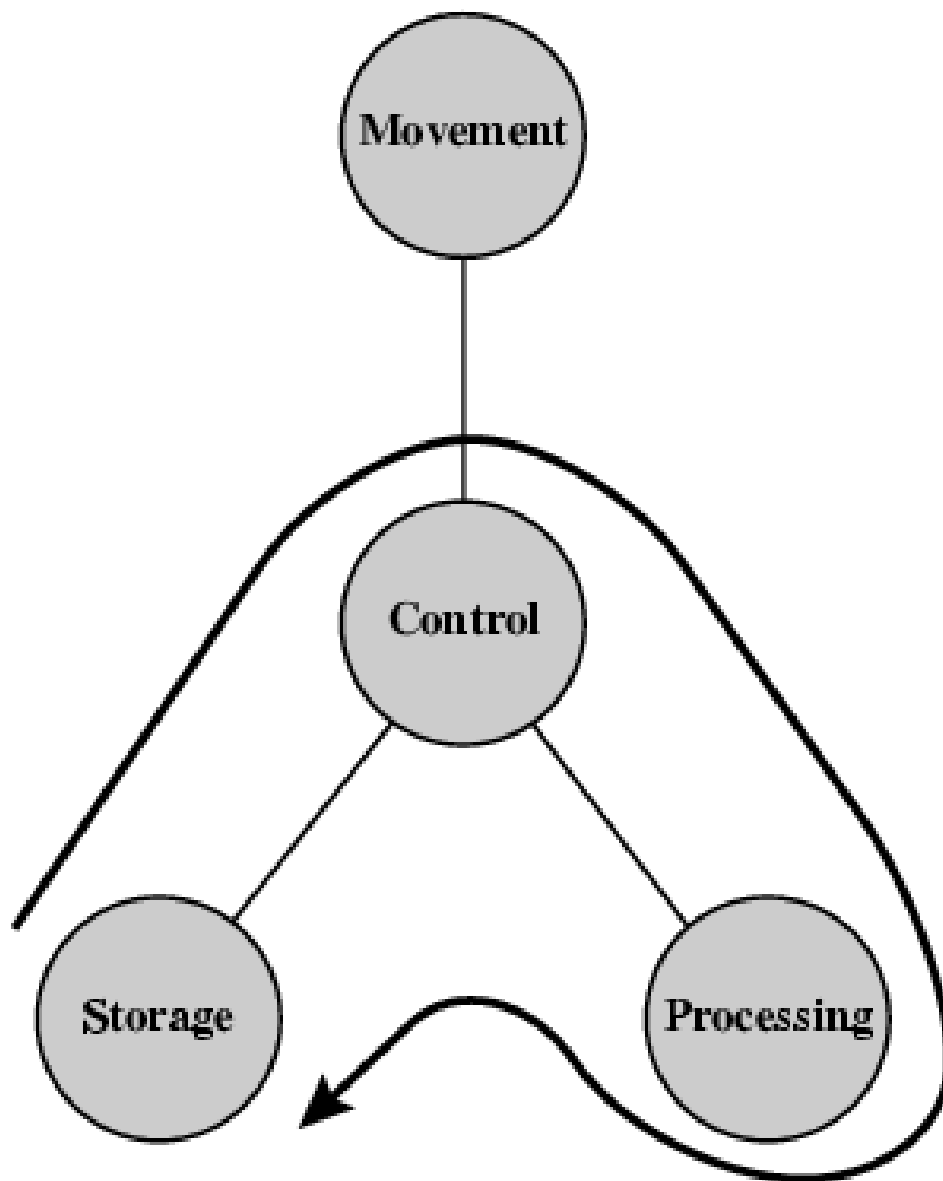
Trao đổi dữ liệu



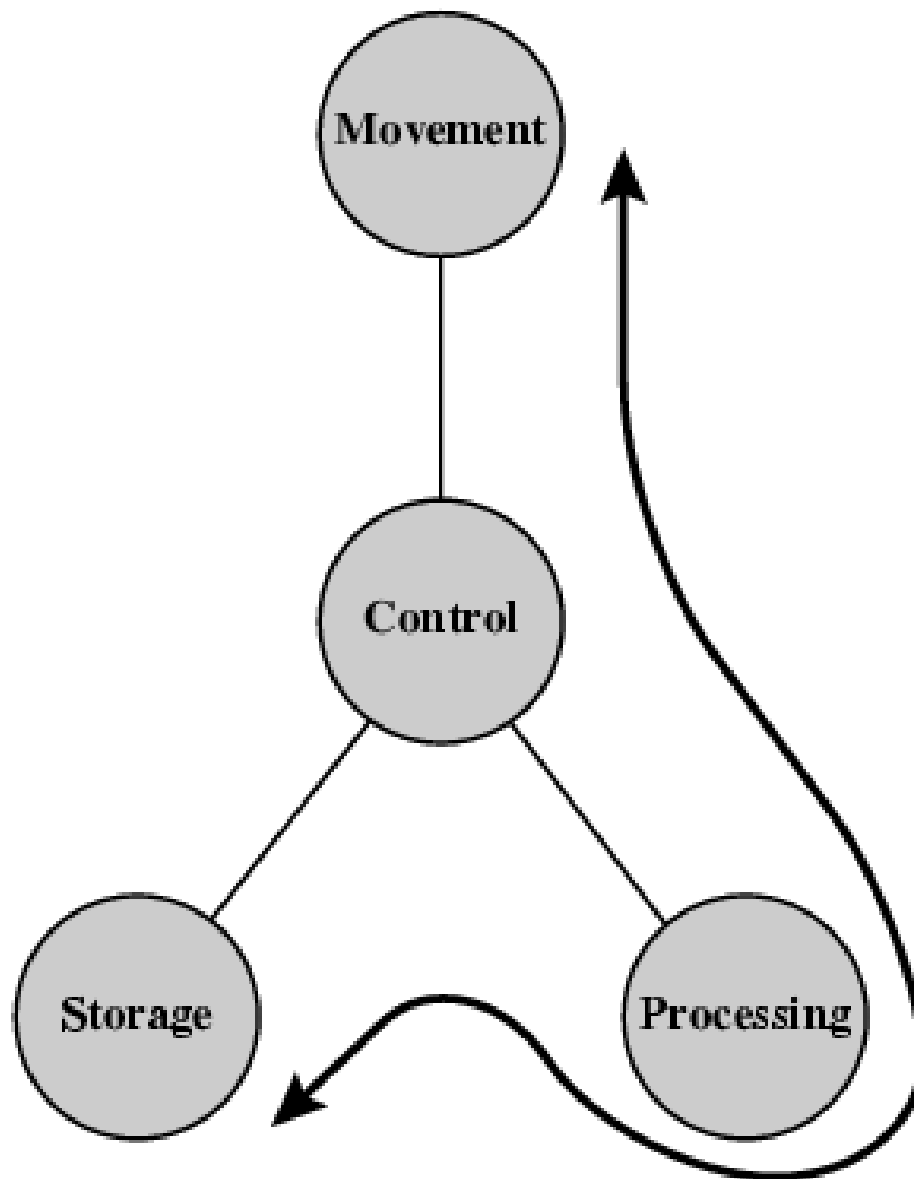
Lưu trữ



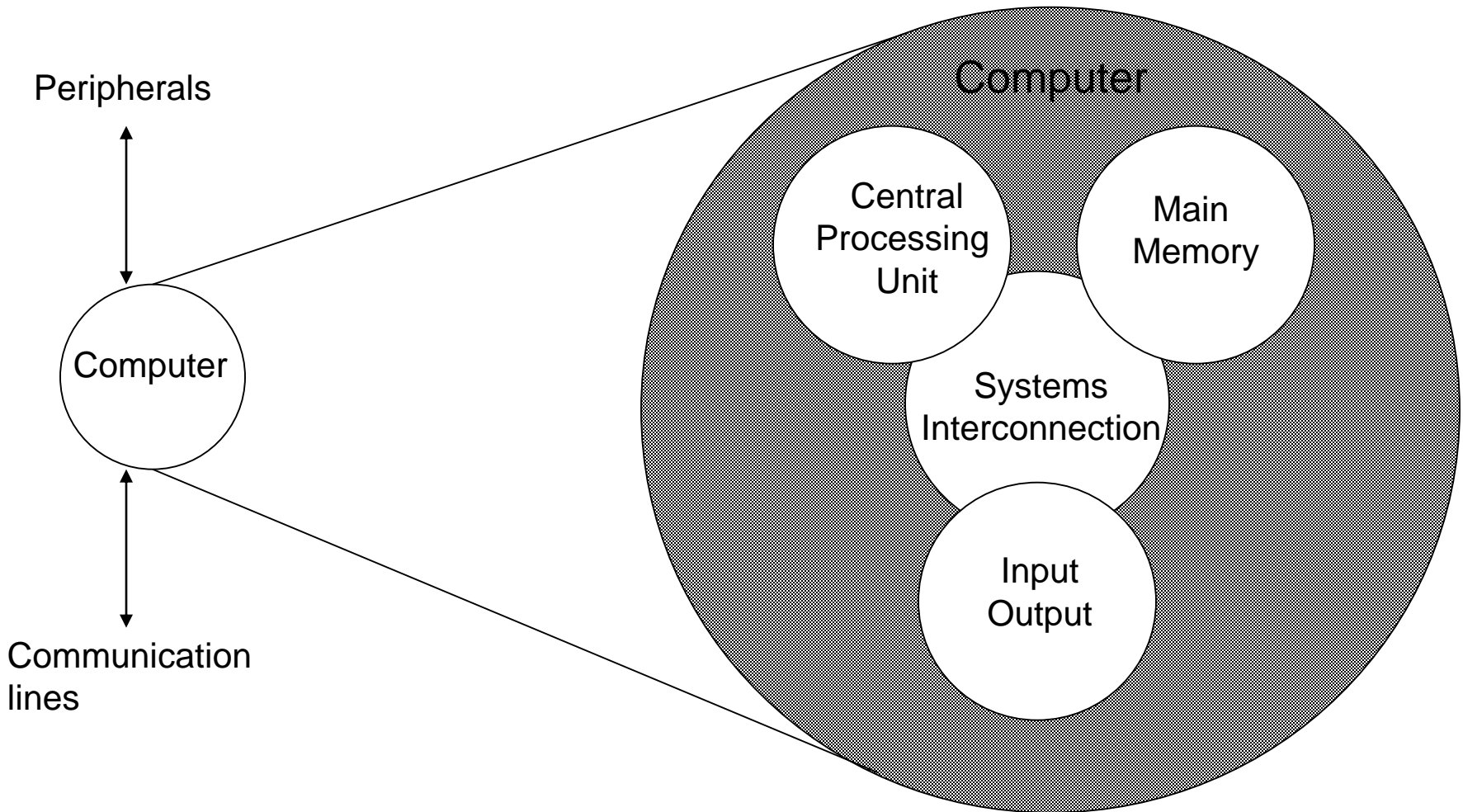
Xử lý thông tin



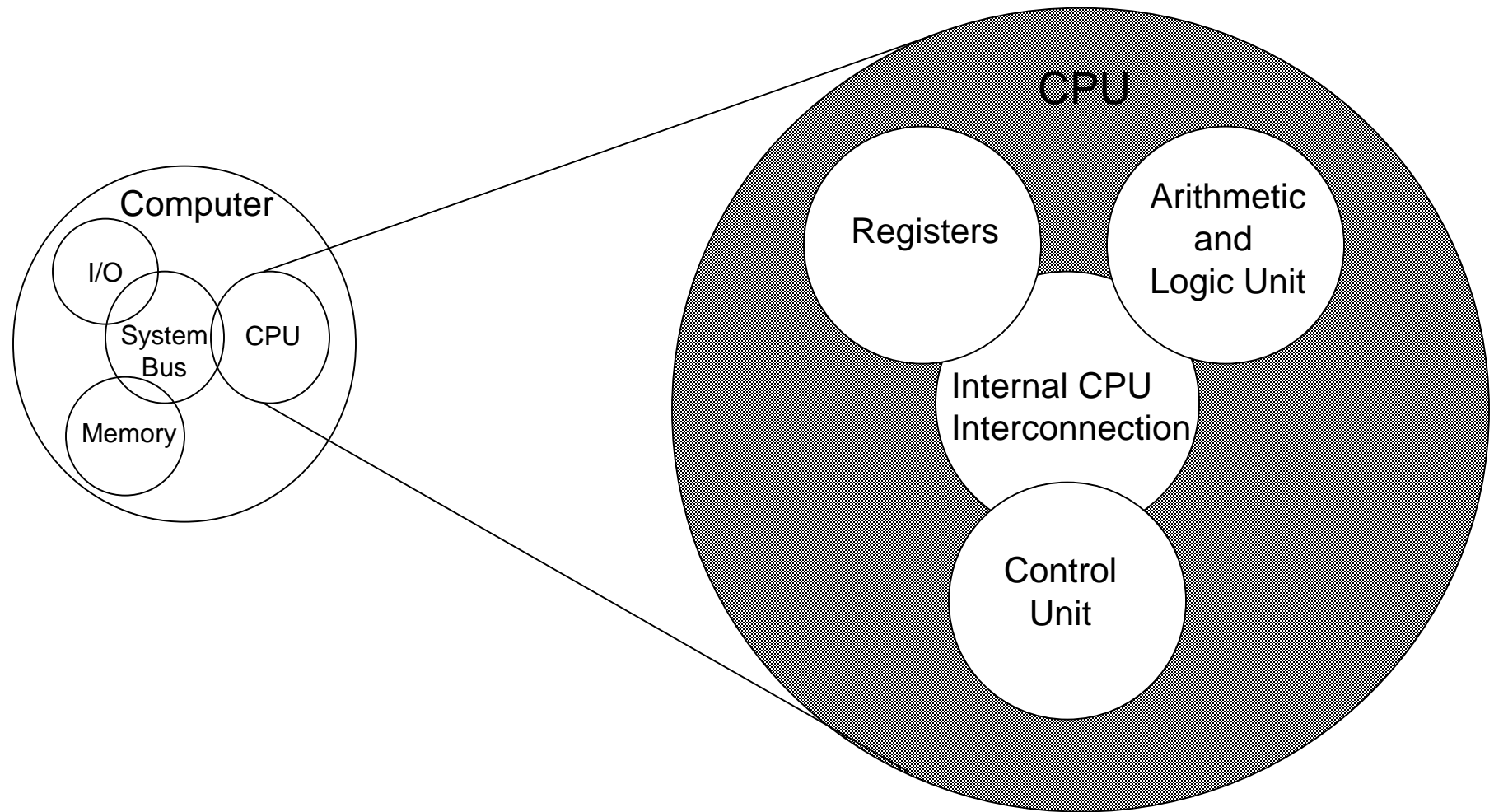
Xử lý và kết xuất thông tin



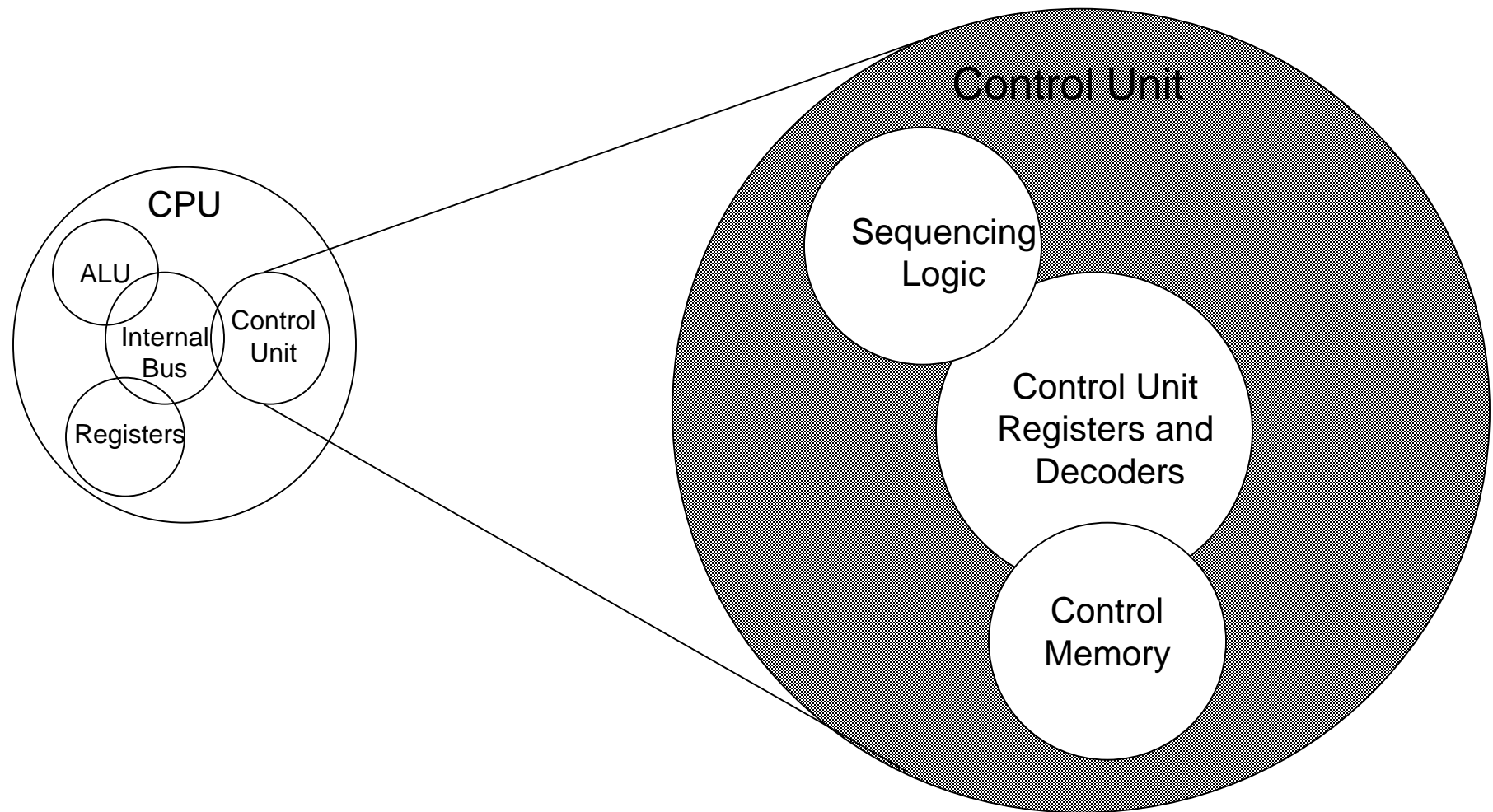
Cấu trúc máy tính – top view



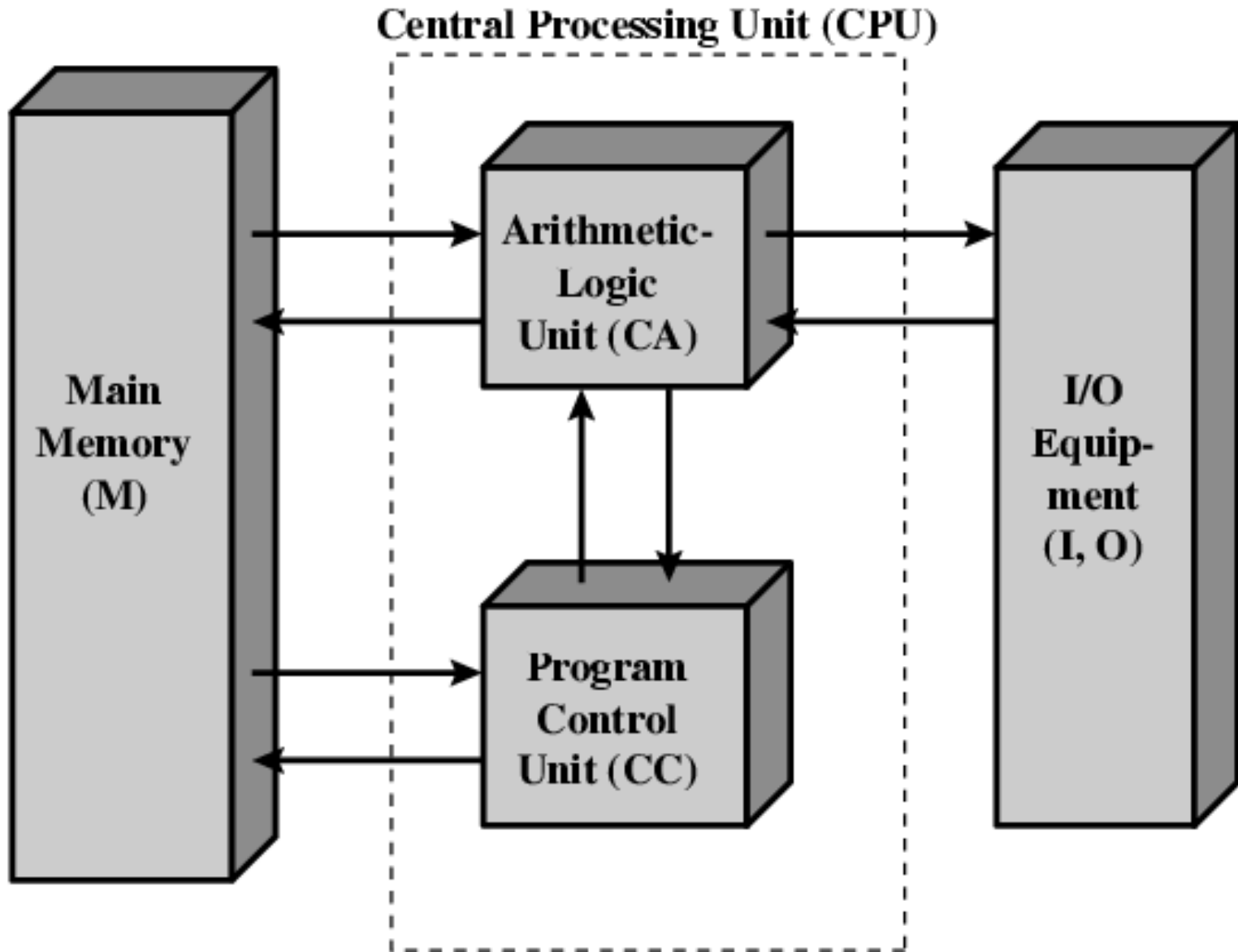
Cấu trúc CPU



Cấu trúc Control Unit

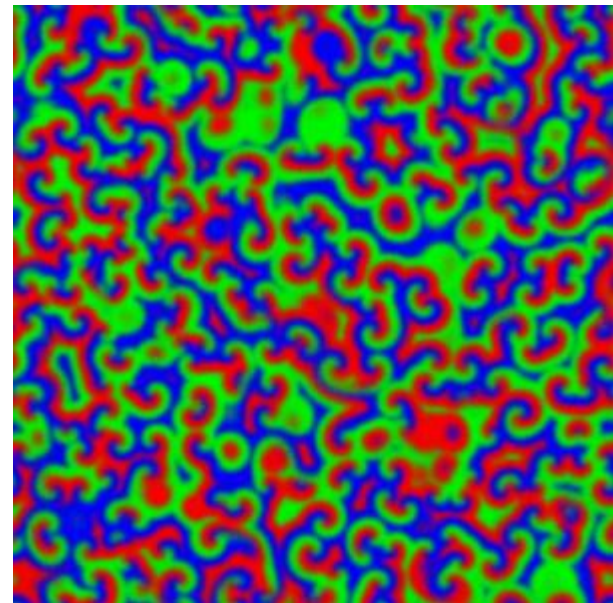
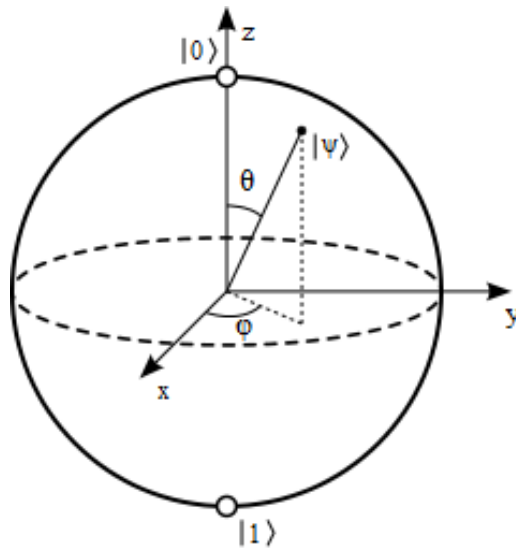


Mô hình máy Von Neuman



Phân loại kiến trúc máy tính

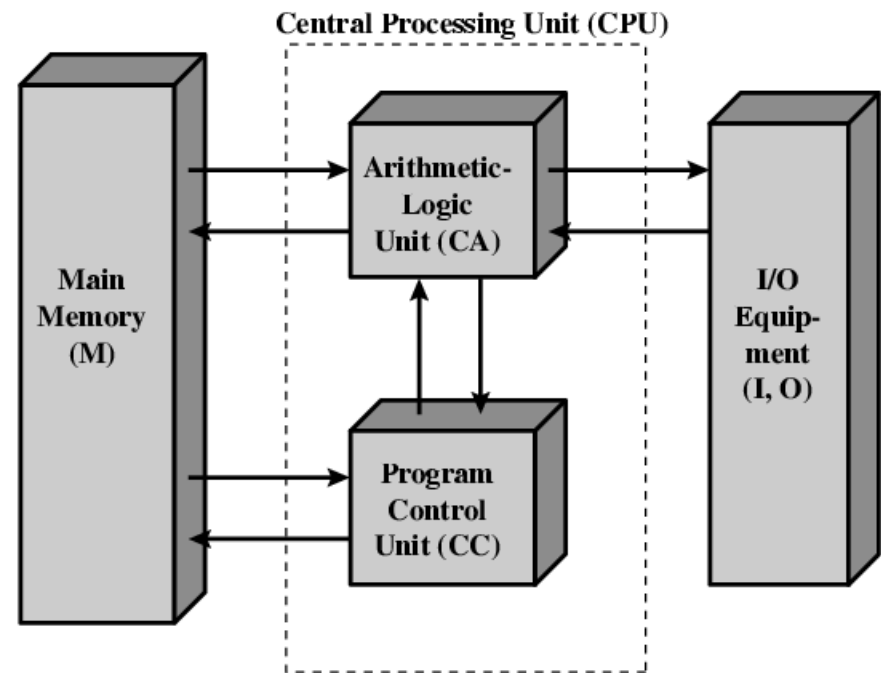
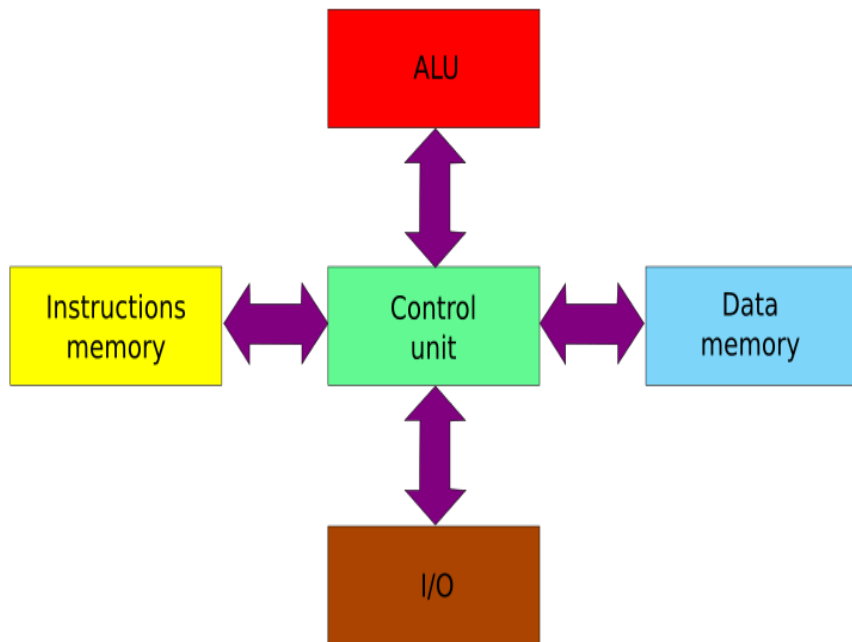
- Scalar Processor <> Vector Processor
- Register machine (Turing) <> Stack machine
- Quantum computer <> Chemical computer



Reaction-diffusion computer
Belousov–Zhabotinsky computer

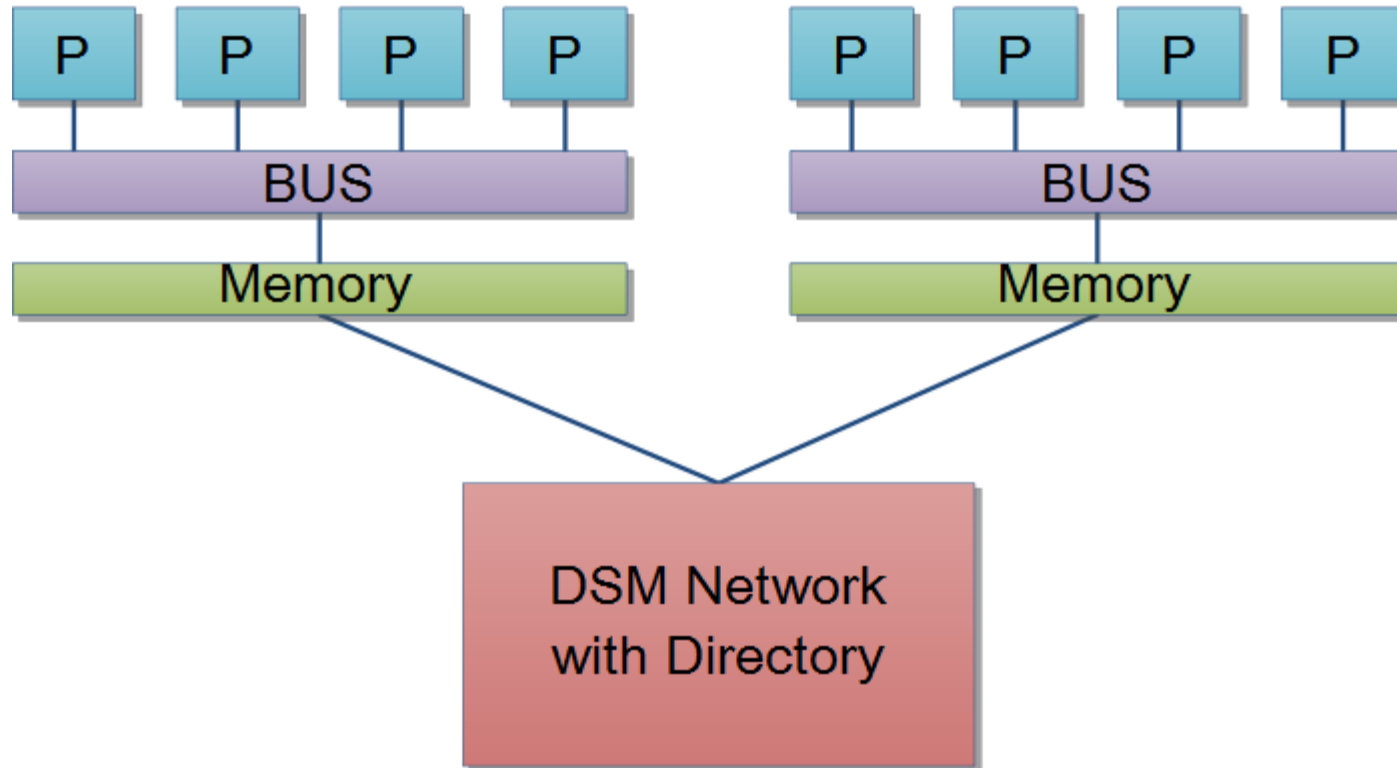
Phân loại kiến trúc máy tính...

- Harvard architecture <> von Neumann architecture



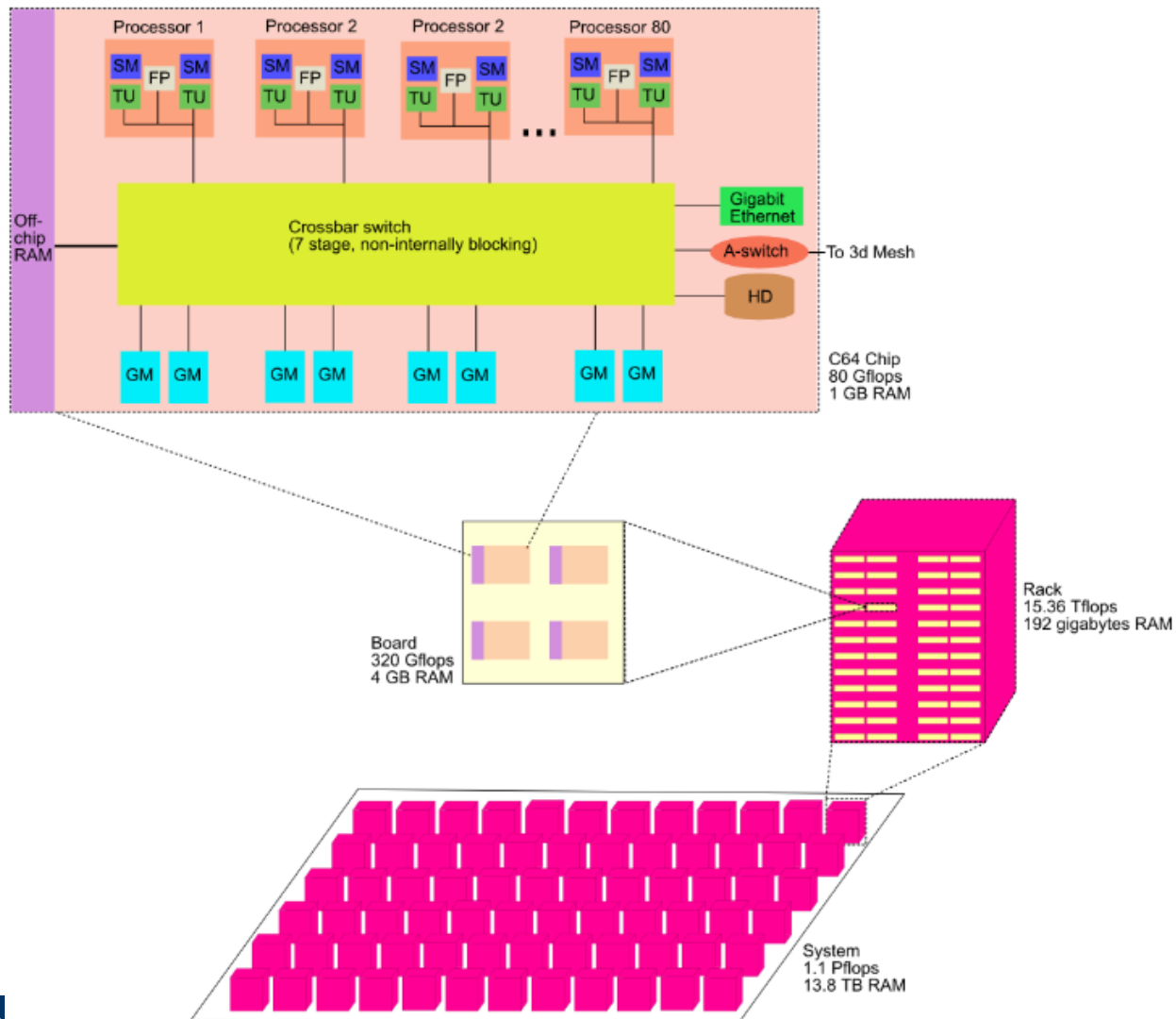
Phân loại kiến trúc máy tính...

- Non-Uniform Memory Access (NUMA) computers



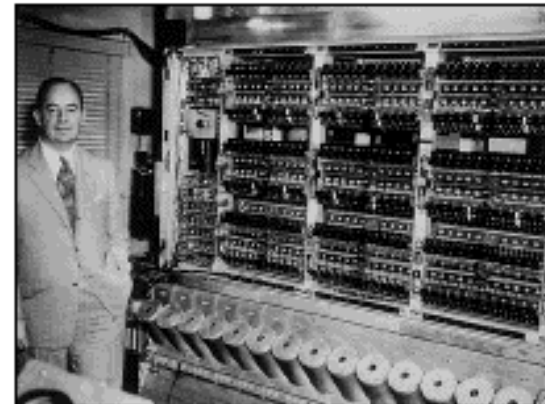
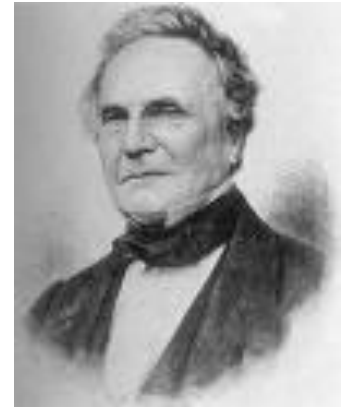
Phân loại kiến trúc máy tính...

- Cellular architecture



Lịch sử

- Trước thế kỷ 19: automates
 - Hộp chơi nhạc
 - Máy dệt
 - ...
- Thế kỷ 19: Charles Babbage đã xây dựng một máy tính – calculator
 - Lập lại những chuỗi thao tác
 - Chọn trạng thái tính
- Thế kỷ 20:
 - 1946: J. Von Neumann miêu tả một mô hình máy tính vạn năng
 - Máy tính hiện nay:
 - Dựa trên mô hình đó
 - Được xếp vào thế hệ thứ 5

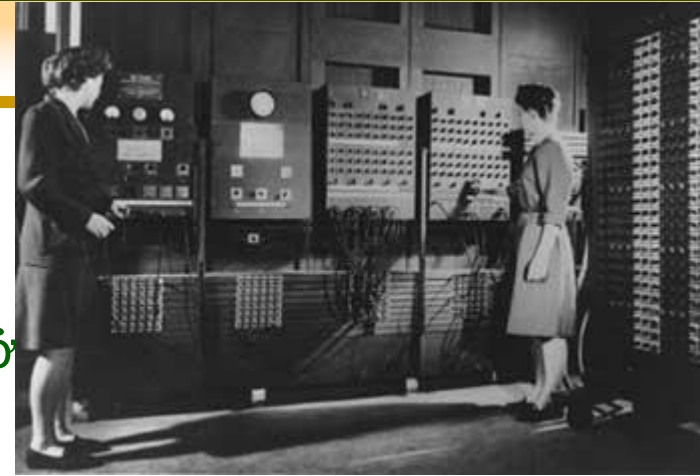


Thế hệ

	Thời gian	Công nghệ	Tốc độ (operations/s)
1	1946-1957	Bóng chân không	40.000
2	1958-1964	Transistor	200.000
3	1965-1971	SSI/MSI	1.000.000
4	1972-1977	LSI	10.000.000
5	1978-	VLSI	100.000.000
6		ULSI	>100.000.000 devices/chip

1946-1957

- Máy tính chuyên dụng, mẫu duy nhất
 - Công kênh, vận hành kém
 - Công nghệ bóng chân không, rơle, điện trở
 - Lập trình bằng thẻ đục lỗ

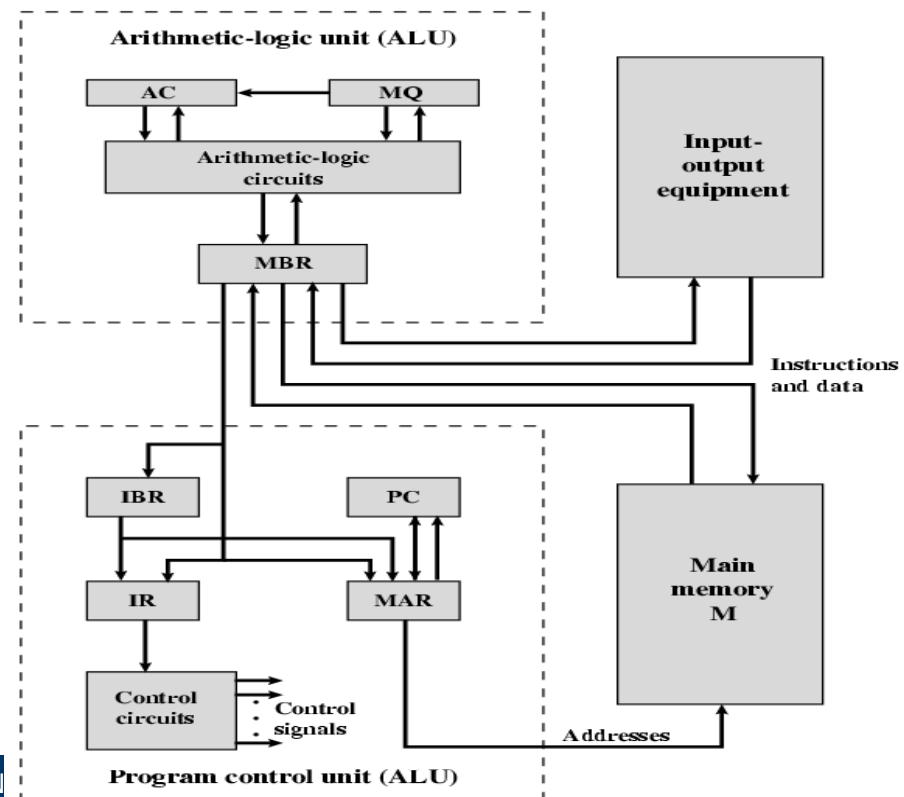
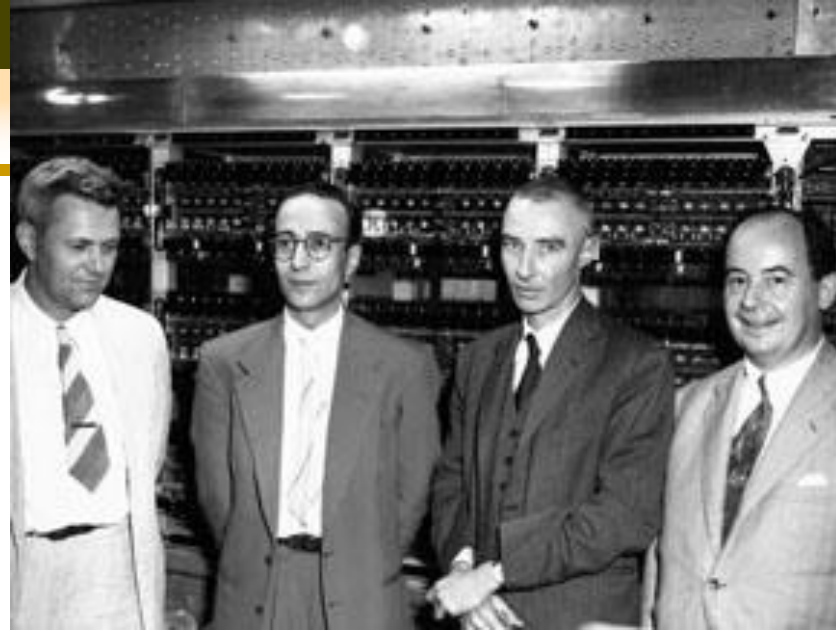


- Tiêu biểu: ENIAC – Electronic Numerical Integrator And Calculator dành để nghiên cứu về bom H
 - 1943-1946, được sử dụng đến 1955
 - 18.000 đèn điện tử, 1.500 rơle, 30 tấn, 140KW, 15.000 square feet
 - 5.000 phép cộng/s
 - 20 thanh ghi, mỗi thanh chứa được 1 số thập phân 10 chữ số



IAS

- Princeton Institute for Advanced Studies, hoàn thành 1952
 - 1000 x 40 bit words
 - Binary number
 - 2 x 20 bit instructions
 - Set of registers (storage in CPU)
 - Memory Buffer Register
 - Memory Address Register
 - Instruction Register
 - Instruction Buffer Register
 - Program Counter
 - Accumulator
 - Multiplier Quotient



1958-1964

- Sử dụng Transistor
 - Transistor: John Barden, Walter Brattain và William Shockley sáng chế ra ở Bell Lab, 1947
- Ngôn ngữ lập trình đầu tiên : COBOL, FORTRAN, LISP
- Tiêu biểu: máy tính mini DEC PDP-1 (1957)
 - 4K từ 18bit
 - Chu kỳ 5 μ s



1965-1971

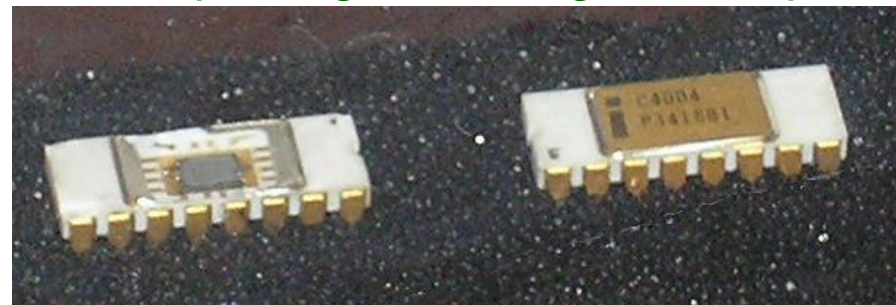
- Công nghệ mạch tích hợp
 - S/MSI : Small/Medium Scale Integration

DEC PDP-8



1972-1977

- Công nghệ LSI (large SI)
 - 10^7 phần tử logics
 - 3.000-100.000 devices/chip
- Mạng máy tính lên ngôi
- Xử lý phân tán
- 1971 : microprocessor 4004 đầu tiên của INTEL
 - Tất các các thành phần của CPU được tập trung trên cùng một chip



1978-

- Công nghệ VL/WSI (very large, wafer)
 - $> 10^8$ phần tử logics
 - 100.000-100.000.000 devices/chip
- Hệ phân tán tương tác
 - Multimedia,
 - Xử lý dữ liệu không phải dạng số (text, images, speech)
 - Song song,
 - Client-server
- Kiểu:
 - MicroComputer – PC
 - MiniComputer
 - SuperMini
 - MainFrame
 - SuperComputer
- Tham khảo: <http://www.computerhistory.org/>

Một số siêu máy tính

	System	Vendor	Total Cores	Rmax (TFlops)	Rpeak (TFlops)	Power (kW)	country
Tianhe-2 (MilkyWay-2)	TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P	NUDT	3,120,000	33,862.70	54,902.40	17,808.00	National University of Defense Technology, CN
Titan - Cray XK7	Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x	Cray Inc.	560,640	17,590.00	27,112.50	8,209.00	US, DOE/SC/Oak Ridge National Laboratory
Sequoia - BlueGene/Q	BlueGene/Q, Power BQC 16C 1.60 GHz, Custom	IBM	1,572,864	17,173.20	20,132.70	7,890.00	US, DOE/NNSA/LLNL
K computer, SPARC64	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect	Fujitsu	705,024	10,510.00	11,280.40	12,659.89	JP, RIKEN Advanced Institute for Computational Science (AICS)
Mira - BlueGene/Q	BlueGene/Q, Power BQC 16C 1.60GHz, Custom	IBM	786,432	8,586.60	10,066.30	3,945.00	US, DOE/SC/Argonne National Laboratory
Stampede - PowerEdge	PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P	Dell	462,462	5,168.10	8,520.10	4,510.00	US, Texas Advanced Computing Center/Univ. of Texas
JUQUEEN - BlueGene/Q	BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	IBM	458,752	5,008.90	5,872.00	2,301.00	Germany, Forschungszentrum Juelich (FZJ)
Vulcan - BlueGene/Q	BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect	IBM	393,216	4,293.30	5,033.20	1,972.00	US, DOE/NNSA/LLNL
SuperMUC - iDataPlex	iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR	IBM	147,456	2,897.00	3,185.10	3,422.67	Germany, Leibniz Rechenzentrum
Tianhe-1A	NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050	NUDT	186,368	2,566.00	4,701.00	4,040.00	National Supercomputing Center in Tianjin

Một số siêu máy tính...



■ Sequoia

- Bộ năng lượng Mỹ với 1,57 triệu nhân xử lý và đạt tốc độ 16,32 petaflop (16,32 triệu tỷ phép tính mỗi giây)
- IBM sản xuất
- Hai hệ thống nữa của Mỹ trong Top 10 Supercomputer 2012 là "báo gấm" Cray Jaguar (đứng thứ 6) và FERMI (thứ 7).

Một số siêu máy tính...



- Fujitsu K
 - 2011, có 705.024 lõi và đạt 10,51 petaflop.
 - 830 megaflop mỗi watt

Một số siêu máy tính...



- SuperMUC, Đức không chỉ ít tốn năng lượng hơn Tianhe-1a mà còn đánh bại cỗ máy tính của Trung Quốc về hiệu suất với tốc độ 2,9 petaflop. Nó phục vụ cho công việc nghiên cứu và giảng dạy tại một số trường đại học ở Munich. Ngoài ra, Đức còn sở hữu JuQueen (xếp thứ 8) dù hệ thống do IBM sản xuất.

Một số siêu máy tính...



- 2010, Tianhe-1a đạt 2,57 petaflop, chiếm ngai vàng mà Mỹ lúc đó nắm giữ. Tuy nhiên, nó khá ngốn năng lượng vì chỉ thực hiện được 635 megaflop mỗi watt. Hệ thống tiếp theo là Nebulae 1,27 petaflop
- http://www.china.org.cn/top10/2013-06/21/content_29187340_10.htm

Một số siêu máy tính...



- Pháp: Curie tốc độ 1,36 petaflop giúp Pháp được nhắc đến trong danh sách 10 siêu máy tính.

Hiệu năng – Performance

- Đánh giá hiệu năng máy tính dựa chủ yếu vào 2 tham số
 - Latency (how long to do X)
 - Còn được gọi là thời gian thực thi / thời gian đáp ứng
 - Throughput (how often can it do X)
- Hiệu năng phụ thuộc vào kiến trúc
 - Tần số xung đồng hồ
- Với CPU, hay sử dụng tham số số lệnh thực hiện được trong 1 giây
 - MIPS : Millions Instructions Per Second
 - MFLOPS : Millions FLOating Point Instructions Per Second
- Hiệu năng toàn hệ thống được xác định thông qua những chương trình đánh giá chuyên biệt benchmarks

Hiệu năng...

■ CPI & IPC

- Thời gian thi hành $T_{\text{exe}} = \text{NI} * \text{CPI} * T_C$

NI: Number of instructions

CPI: Cycle per Instruction

T_C : Hàm công nghệ cho phép xác định thời gian thi hành các bước cơ bản

- $\text{IPC} = 1/\text{CPI}$: Instructions per Cycle

- $N_{\text{MIPS}} = \text{NI} \times 10^{-6} / T_{\text{exe}} = F \times \text{IPC}$ (F tần số được đo bằng MHz)

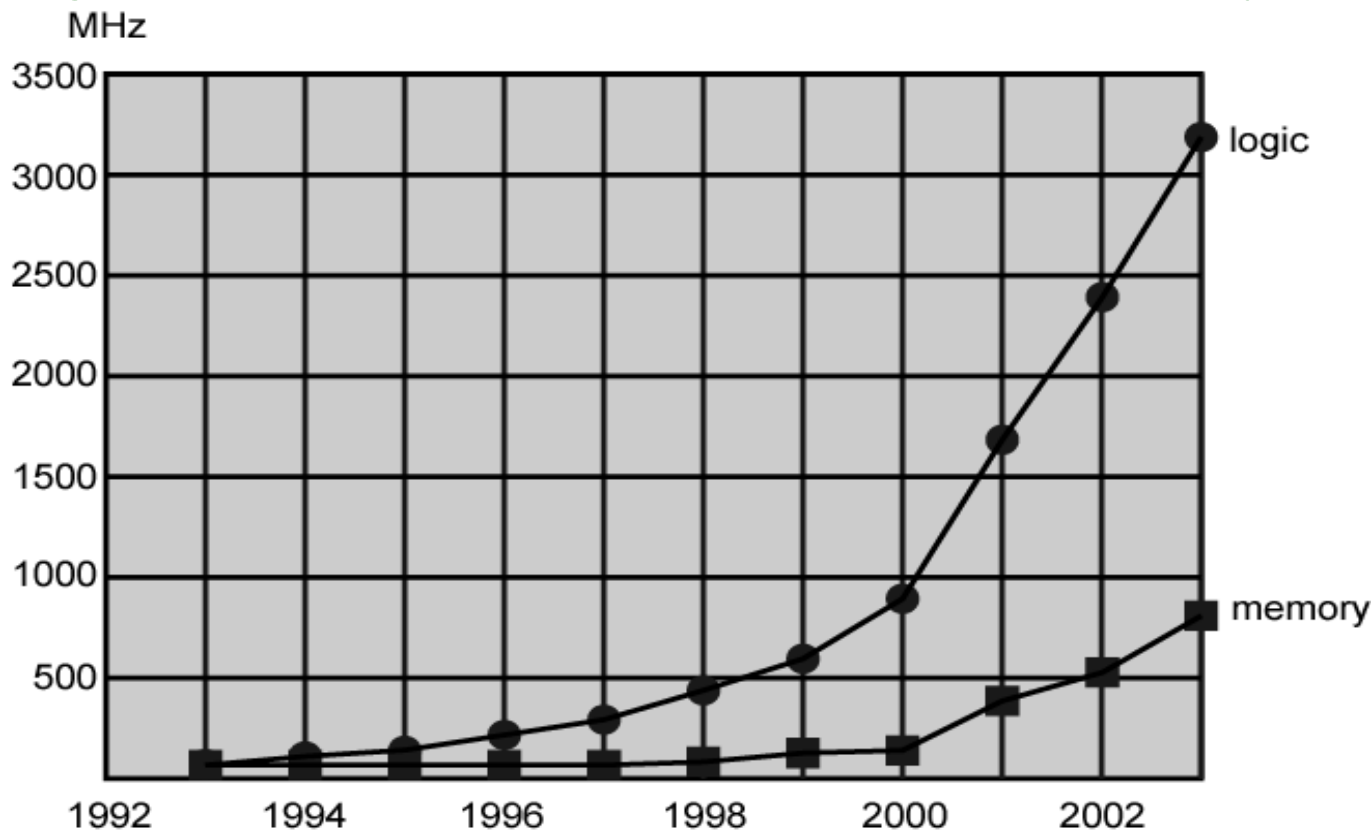
- CPI thể hiện hiệu năng của các kiến trúc cứng: cho phép so sánh hiệu năng của các phần cứng được thực thi từ cùng một kiến trúc mềm,
- CPI cho phép so sánh các kiến trúc của bộ vi xử lý (chẳng hạn RISC và CISC).

Luật Moore - 1965

- Số transistors trên một đơn vị diện tích sẽ tăng gấp đôi sau mỗi 24 tháng
 - <1970, gấp đôi mỗi năm
 - David House: hiệu năng sẽ tăng gấp đôi sau mỗi 18 tháng
- Giá thành chip không đổi
- Mật độ cao hơn → hiệu năng cao hơn
- Kích thước bé hơn → mức độ linh hoạt cao hơn (flexibility)
- Giảm công suất tiêu thụ và cần có các thiết bị làm lạnh
- Giảm hệ thống liên kết giữa các phần tử, tăng độ tin cậy của chip
- Một số khuynh hướng dẫn xuất:
 - Processor performance : Twice as fast after ~18 months
 - Memory capacity : Twice as much in <2 years
- <http://www.intel.com/research/silicon/mooreslaw.htm>

Các yếu tố tác động đến hiệu năng

- Tốc độ bộ vi xử lý
- Dung lượng bộ nhớ
 - Ràng buộc: tốc độ bộ nhớ luôn đi sau tốc độ bộ vi xử lý



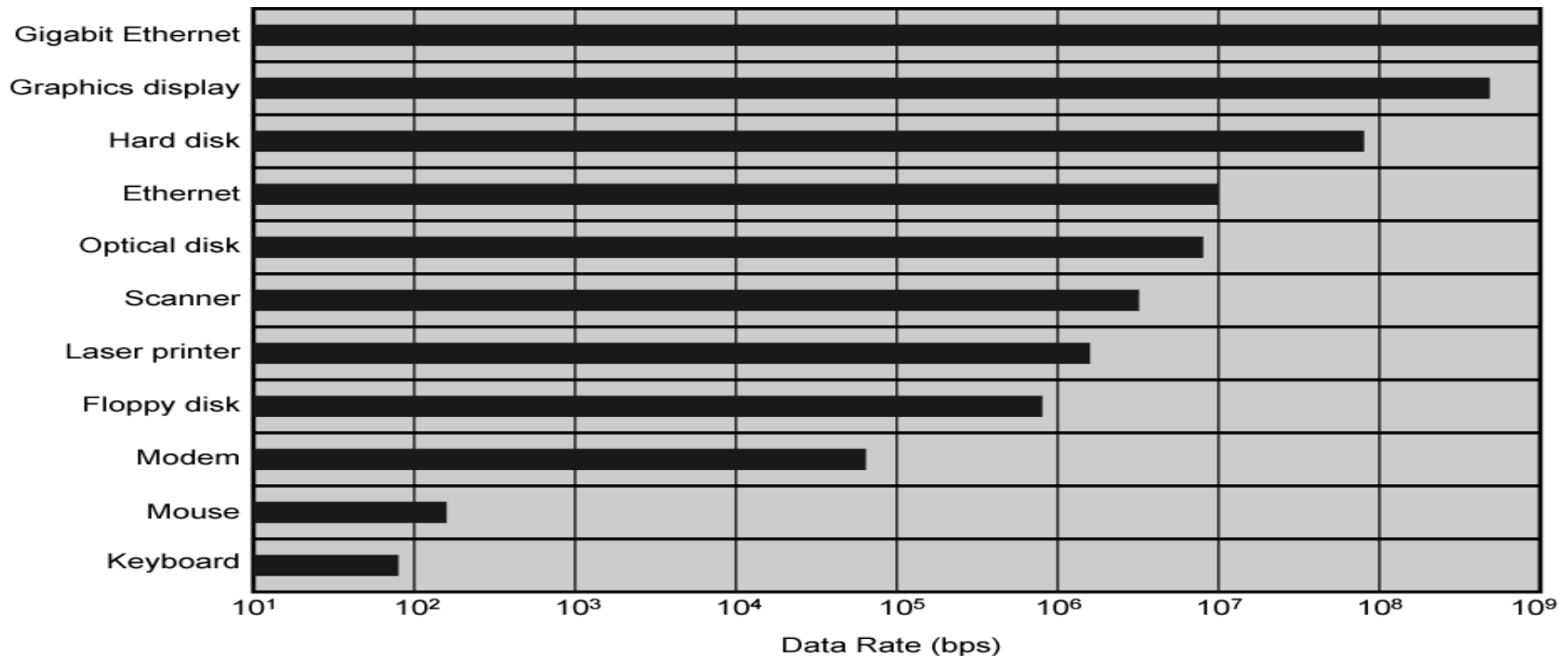
- Hiệu năng truy cập dữ liệu từ thiết bị vào/ra

Giải pháp tăng hiệu năng

- Tăng hiệu năng các thành phần cấu thành máy tính
- Tăng số bits thao tác tại mỗi thời điểm (8-16-32-64, ...)
 - Tạo DRAM “rộng hơn” thay vì “nhanh hơn”
- Thay đổi tổ chức/kiến trúc bộ nhớ
 - Phân cấp bộ nhớ, sử dụng bộ nhớ đệm cache
- Giảm tần xuất truy cập bộ nhớ
 - Cache phức tạp hơn và cache ngay trong chip
- Tăng thông lượng liên kết
 - Bus tốc độ cao
 - Phân cấp bus

Giải pháp...

- Đối với các thiết bị ngoại vi (I/O)
 - Caching
 - Buffering
 - Higher-speed interconnection buses
 - More elaborate bus structures
 - Multiple-processor configurations



Thực tế

- Cân đối những yếu tố quyết định đến hiệu năng
 - Processor components
 - Main memory
 - I/O devices
 - Interconnection structures

Thực tế ...

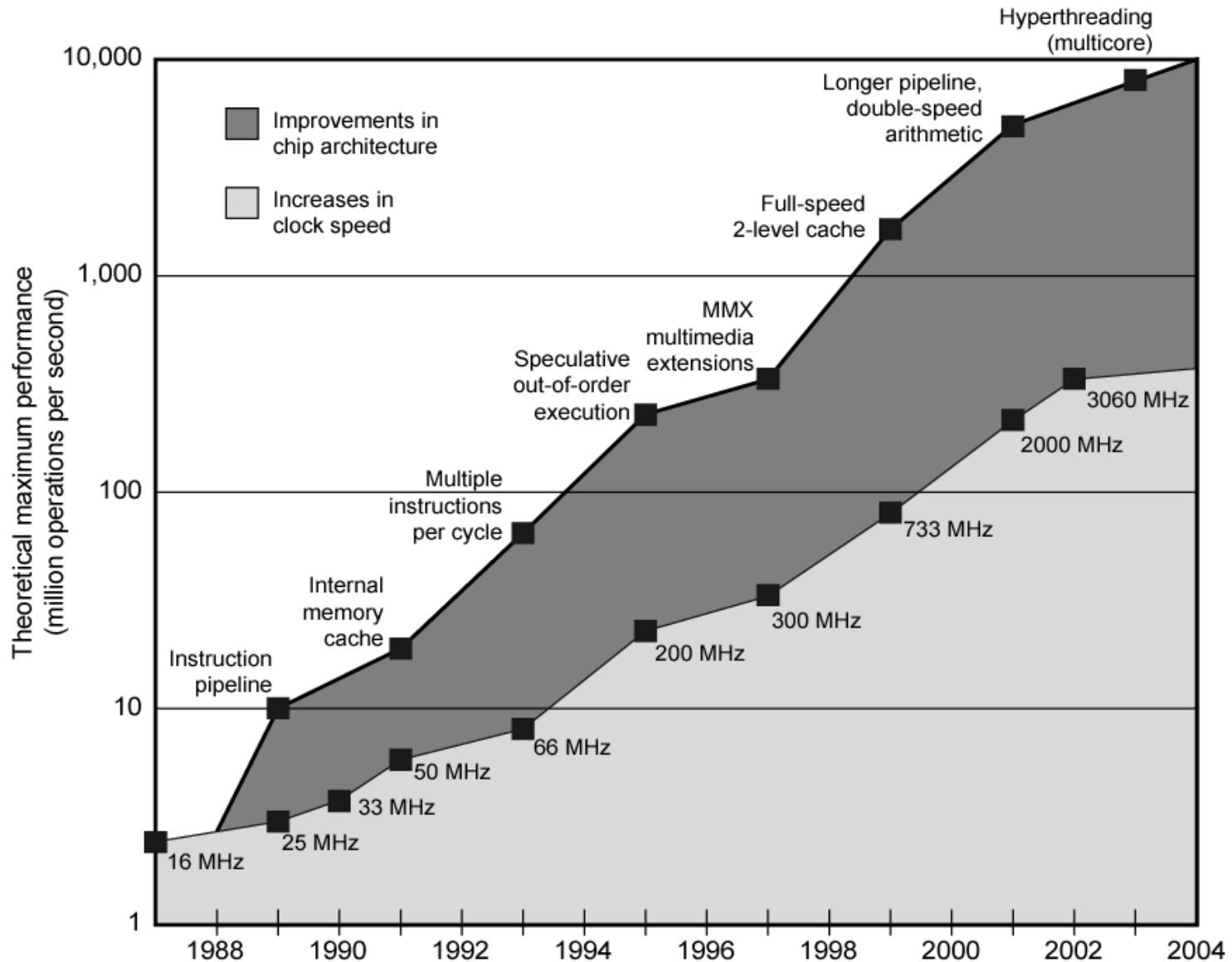
- Tăng tốc độ phần cứng của vi xử lý
 - Giảm kích thước các cổng logic
 - Tăng xung nhịp đồng hồ
 - Giảm thời gian lan truyền tín hiệu giữa các cổng logic
- Tăng dung lượng và tốc độ caches
 - Đưa một phần cache vào bên trong chip → giảm thời gian truy cập cache
 - Phân cấp caches : thường có 2-3 mức cache giữa bộ nhớ và bộ vi xử lý
 - Tăng mật độ nhớ
- Thay đổi tổ chức và kiến trúc vi xử lý
 - Tăng tốc độ thực thi
 - Xong xong hoá

Tuy nhiên...

Hiệu năng còn phụ thuộc vào

- Xung đồng hồ
 - Mật độ các cổng logics
 - Điện năng
 - Tăng cùng với mật độ các cổng logics và tốc độ xung clock
 - Mất nhiệt năng
 - Độ trễ trở/dung kháng - RC delay
 - Tốc độ dòng electrons giới hạn bởi trở kháng/dung kháng của vật liệu sử dụng
 - Độ trễ tăng khi trở kháng tăng
 - Mạng liên kết càng bé, trở kháng càng tăng
 - Các cổng logics càng gần nhau, dung kháng càng tăng
 - Memory latency
 - Tốc độ bộ nhớ cản trở tốc độ bộ vi xử lý
- Giải pháp:
- Các cách tiếp cận mới về tổ chức và kiến trúc máy tính

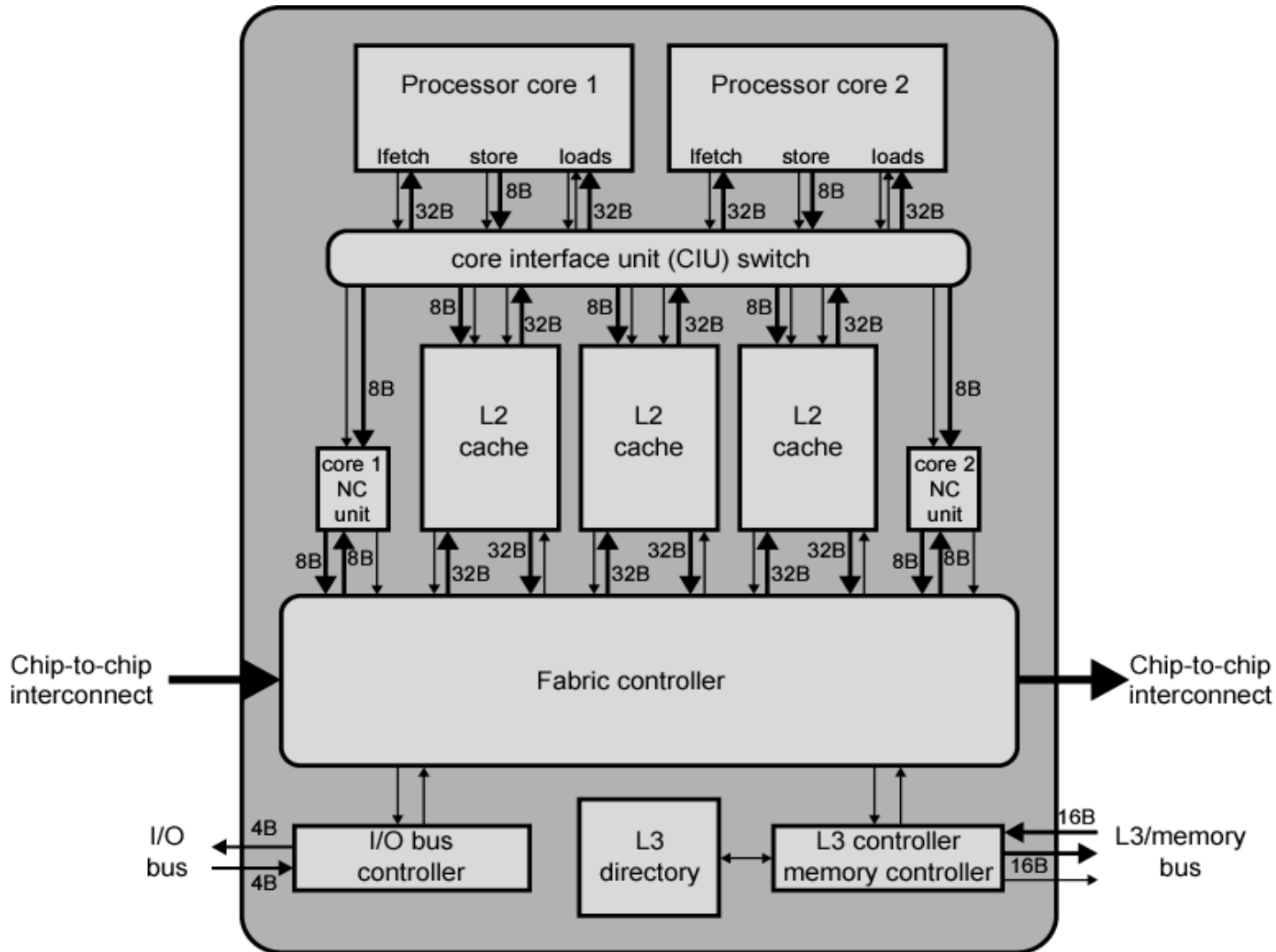
Intel Microprocessor Performance



Đa lõi - Multiple Cores

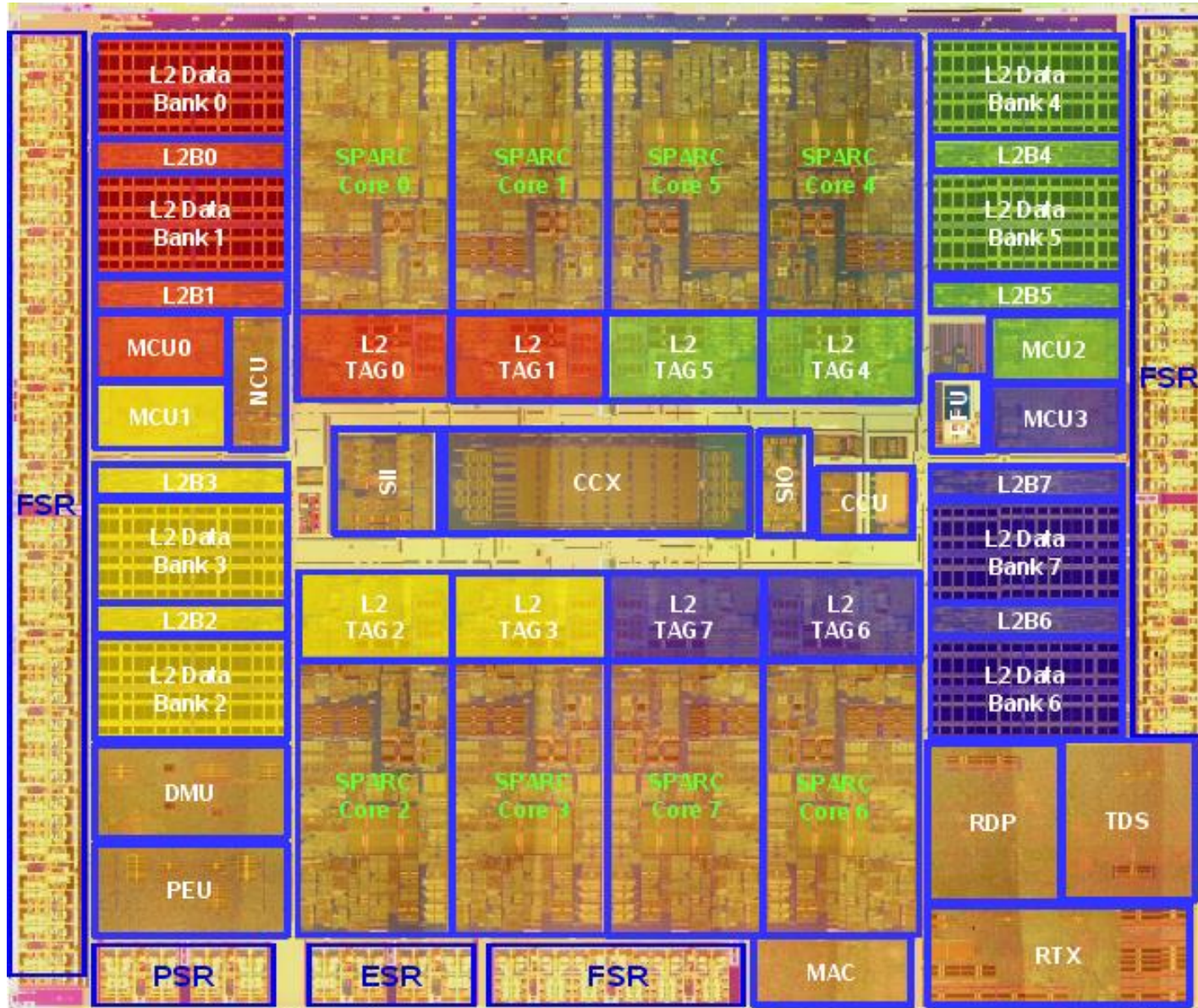
- Kiến trúc đa lõi trên cùng một chip
 - Chia sẻ bộ nhớ cache lớn
- Why ?
 - Nếu phần mềm có thể tận dụng đa VXL, tăng đôi VXL đồng nghĩa với khả năng tăng đôi hiệu năng
 - Với nhiều lõi, bộ nhớ cache lớn phát huy tốt ưu điểm
 - Mức tiêu thụ năng lượng của các phần tử nhớ thấp hơn các phần tử xử lý logic
- Ví dụ:
 - IBM POWER4: two cores based on PowerPC
 - INTEL: DualCore, Core 2 Duo, QuadCore, ...

POWER4 Chip Organization



NC = noncacheable

UltraSparc T2



Intel Processor Evolution (1)

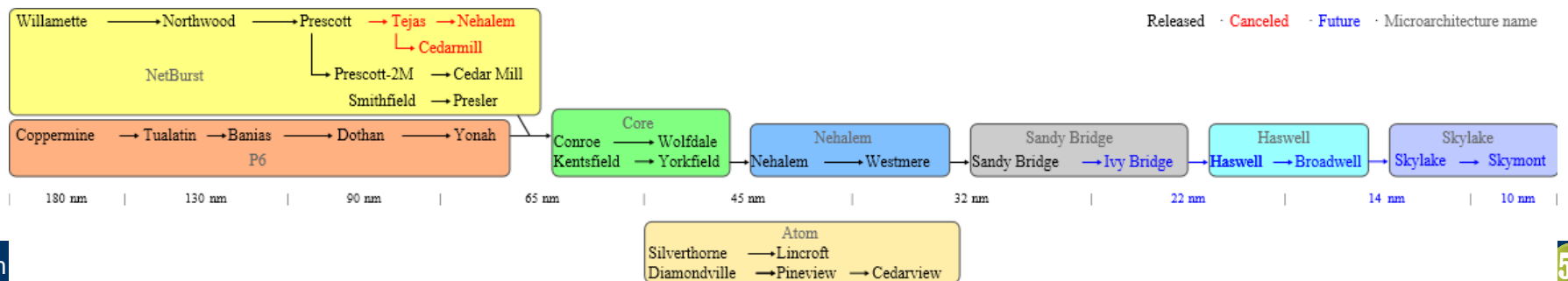
- 4-bit Processor
 - Intel 4004, 4040
- 8-bit Processor
 - 8080
 - first general purpose microprocessor
 - 8 bit data path
 - Used in first personal computer – Altair
- 16-bit processors: origin of x86
 - 8086: much more powerful, instruction cache, prefetch few instructions
 - 8088 (8 bit external bus) used in first IBM PC
 - 80286: 16 Mbyte memory addressable, up from 1Mb
- 32-bit processors:
 - 80386: Support for multitasking
 - 80486: sophisticated powerful cache and instruction pipelining, built in maths co-processor

Intel Processor Evolution(2)

- 32-bit processors: P5 microarchitecture
 - Pentium: Superscalar, Multiple instructions executed in parallel
 - Pentium with MMX technology
- 32-bit processors: P6/Pentium M microarchitecture
 - Pentium Pro: Increased superscalar organization, Aggressive register renaming, branch prediction, data flow analysis, speculative execution
 - Pentium II: MMX technology, graphics, video & audio processing
 - Celeron
 - Pentium III: Additional floating point instructions for 3D graphics
 - Pentium II and III Xeon
 - Pentium M
 - Celeron M
 - Intel Core
 - Dual-Core Xeon LV
- 32-bit processors: NetBurst microarchitecture
 - Pentium 4: Further floating point and multimedia enhancements
 - Xeon
 - Mobile Pentium 4-M
 - Pentium 4 EE, 4E, 4F
- 32-bit processors: Intel 32 – Intel Atom

Intel Processor Evolution (3)

- 64-bit processors: IA-64
 - Itanium 1, 2 : Hardware enhancements to increase speed
- 64-bit processors: Intel 64 – NetBurst microarchitecture
 - Pentium 4F, Pentium D
 - Pentium Extreme Edition
 - Xeon
- 64-bit processors: Intel 64 – Core microarchitecture
 - Xeon
 - Intel Core 2
 - Pentium Dual Core (Wolfdale-3M 45 nm)
 - Celeron, Celeron M
- 64-bit processors: Intel 64 – Nehalem microarchitecture
 - Intel Pentium (Clarkdale – 32 nm process technology)
 - Core i3, i5, i7, Xeon
- 64-bit processors: Sandy Bridge



PowerPC

- 1975, 801 minicomputer project (IBM) RISC
- Berkeley RISC I processor
- 1986, IBM commercial RISC workstation product, RT PC
 - Not commercial success
 - Many rivals with comparable or better performance
- 1990, IBM RISC System/6000
 - RISC-like superscalar machine
 - POWER architecture
- IBM alliance with Motorola (68000 microprocessors), and Apple, (used 68000 in Macintosh)
- Result is PowerPC architecture
 - Derived from the POWER architecture
 - Superscalar RISC
 - Apple Macintosh
 - Embedded chip applications



PowerPC Family (1)

- 601:
 - Quickly to market. 32-bit machine
- 603:
 - Low-end desktop and portable
 - 32-bit
 - Comparable performance with 601
 - Lower cost and more efficient implementation
- 604:
 - Desktop and low-end servers
 - 32-bit machine
 - Much more advanced superscalar design
 - Greater performance
- 620:
 - High-end servers
 - 64-bit architecture

PowerPC Family (2)

- 740/750:
 - Also known as G3
 - Two levels of cache on chip
- G4 (7xxx - 1999):
 - Increases parallelism and internal speed
- 970 - G5 (2003):
 - Improvements in parallelism and internal speed
 - 64-bit organization
- Current: e200, e300, e500, e600, e5500, POWER6, POWER7, PPC4xx, PPC7xx, PPC A2, Xenon (Xbox), Cell(PS3), Broadway...

Tổng kết

- Một số khái niệm cơ bản
 - Kiến trúc & Tổ chức máy tính
 - Cấu trúc và chức năng
 - Mô hình máy tính Von Neuman
- Sơ lược về quá trình phát triển của máy tính
- Các yếu tố liên quan đến hiệu năng và cách thức nâng cao hiệu năng
- Sơ lược về các dòng chip thông dụng của Intel và IBM

Chuẩn bị cho bài sau: Đọc chương 3 của [1]

Câu hỏi ôn tập

1. Tìm hiểu khái niệm “họ - family” sản phẩm ?
2. Phân biệt khái niệm kiến trúc và tổ chức máy tính ?
3. Phân biệt khái niệm cấu trúc và chức năng các thành phần trong máy tính?
4. Chức năng chính của máy tính là gì? Mối liên hệ giữa các chức năng đó?
5. Nguyên lý cơ bản để chế tạo các máy tính là gì ? Tham số nào thường được sử dụng để phân loại các thế hệ máy tính? Các kiểu máy tính?
6. Hiệu năng máy tính phụ thuộc vào những tham số nào? Làm thế nào để nâng cao hiệu năng máy tính?
7. Một máy tính có tần số xung nhịp chip là 5GHz thi hành 1 chương trình bao gồm 5 tỷ lệnh. Số lệnh này gồm 20% lệnh rẽ nhánh, 10% lệnh store, 20% lệnh load, và 50% lệnh số học+logic (ALU). Chỉ số trung bình IPC là 1 đối với lệnh rẽ nhánh, 0.5 với lệnh load, 1 với lệnh store, và 2 với các lệnh ALU. Hãy tính thời gian thực thi chương trình này?