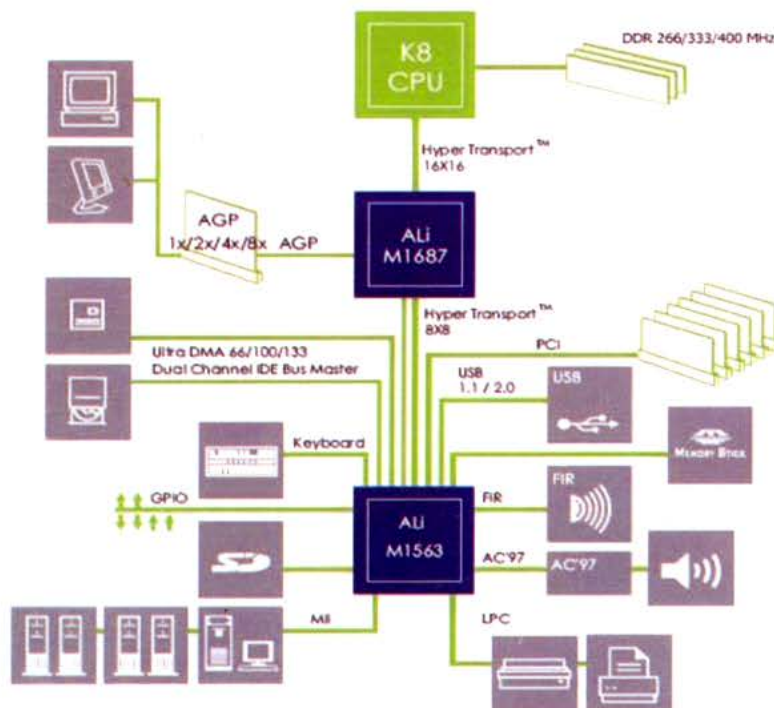


CẤU TRÚC MÁY TÍNH



Biên soạn: NGUYỄN THANH ĐĂNG

(Sử dụng cho hệ Cao đẳng)
Lưu hành nội bộ

NĂM 2009

TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP. HỒ CHÍ MINH
KHOA CÔNG NGHỆ ĐIỆN TỬ

CẤU TRÚC MÁY TÍNH

(Sử dụng cho hệ Cao đẳng)

Biên soạn: NGUYỄN THANH ĐĂNG

**LƯU HÀNH NỘI BỘ
2009**

CHƯƠNG 1

KIẾN TRÚC CƠ BẢN CỦA MÁY TÍNH

§ 1. Những thành phần cơ bản của máy tính

Biểu diễn thông tin trong máy tính

I. Hệ đếm nhị phân và phương pháp biểu diễn thông tin trong máy tính.

1. Hệ nhị phân (Binary)

1.1. Khái niệm:

Hệ nhị phân hay hệ đếm cơ số 2 chỉ có hai con số 0 và 1. Đó là hệ đếm dựa theo vị trí. Giá trị của một số bất kỳ nào đó tùy thuộc vào vị trí của nó. Các vị trí có trọng số bằng bậc lũy thừa của cơ số 2. Chấm cơ số được gọi là chấm nhị phân trong hệ đếm cơ số 2. Mỗi một con số nhị phân được gọi là một bit (Binary digit). Bit ngoài cùng bên trái là bit có trọng số lớn nhất (MSB, Most Significant Bit) và bit ngoài cùng bên phải là bit có trọng số nhỏ nhất (LSB, Least Significant Bit) như dưới đây:

	2^3	2^2	2^1	2^0	2^{-1}	2^{-2}	
MSB	1	0	1	0	.	1	1 LSB
Chấm nhị phân							

Số nhị phân $(1010.11)_2$ có thể biểu diễn thành:

$$(1010.11)_2 = 1*2^3 + 0*2^2 + 1*2^1 + 0*2^0 + 1*2^{-1} + 1*2^{-2} = (10.75)_{10}.$$

Chú ý: dùng dấu ngoặc đơn và chỉ số dưới để ký hiệu cơ số của hệ đếm.

Đối với phần lẻ của các số thập phân, số lẻ được nhân với cơ số và số nhớ được ghi lại làm một số nhị phân. Trong quá trình biến đổi, số nhớ đầu chính là bit MSB và số nhớ cuối là bit LSB.

Ví dụ 2: Biến đổi số thập phân $(0.625)_{10}$ thành nhị phân:

$0.625*2 = 1.250$. Số nhớ là 1, là bit MSB.

$0.250*2 = 0.500$. Số nhớ là 0

$0.500*2 = 1.000$. Số nhớ là 1, là bit LSB.

Vậy : $(0.625)_{10} = (0.101)_2$.

2. Hệ thập lục phân (Hexadecima).

2.1. Khái niệm:

Các hệ máy tính hiện đại thường dùng một hệ đếm khác là hệ thập lục phân.

Hệ thập lục phân là hệ đếm dựa vào vị trí với cơ số là 16. Hệ này dùng các con số từ 0 đến 9 và các ký tự từ A đến F như trong bảng sau:

Bảng 1.1 Hệ thập lục phân

Thập lục phân	Thập phân	Nhi phân
0	0	0000
1	1	0001
2	2	0010
3	3	0011
4	4	0100
5	5	0101
6	6	0110
7	7	0111
8	8	1000
9	9	1001
A	10	1010
B	11	1011
C	12	1100
D	13	1101
E	14	1110
F	15	1111

3. Bảng mã ASCII.(American Standard Code for Information Interchange).

Người ta đã xây dựng bộ mã để biểu diễn cho các ký tự cũng như các con số Và các ký hiệu đặc biệt khác. Các mã đó gọi là bộ mã ký tự và số. Bảng mã ASCII là mã 7 bit được dùng phổ biến trong các hệ máy tính hiện nay. Với mã 7 bit nên có $2^7 = 128$ tổ hợp mã. Mỗi ký tự (chữ hoa và chữ thường) cũng như các con số thập phân từ 0..9 và các ký hiệu đặc biệt khác đều được biểu diễn bằng một mã số như bảng 2-2.

Việc biến đổi thành ASCII và các mã ký tự số khác, tốt nhất là sử dụng mã tương đương trong bảng. Ví dụ: Đổi các ký tự BILL thành mã ASCII:

Ký tự	B	I	L	L
ASCII	1000010	1001001	1001100	1001100
HEXA	42	49	4C	4C

Bảng 1.2: Mã ASCII.

						Column bits(B ₇ B ₆ B ₅)							
Bits(row)						000	001	010	011	100	101	110	111
ROW	B ₄	B ₃	B ₂	B ₁		0	1	2	3	4	5	6	7
						↓	↓	↓	↓	↓	↓	↓	↓
0	0	0	0	0	→	NUL	DLE	SP	0	@	P	\	p
1	0	0	0	1	→	SOH	DC1	!	1	A	Q	A	q
2	0	0	1	0	→	STX	DC2	“	2	B	R	B	r
3	0	0	1	1	→	ETX	DC3	#	3	C	S	C	s
4	0	1	0	0	→	EOT	DC4	\$	4	D	T	D	t
5	0	1	0	1	→	ENQ	NAK	%	5	E	U	E	u
6	0	1	1	0	→	ACK	SYN	&	6	F	V	F	v
7	0	1	1	1	→	BEL	ETB	‘	7	G	W	G	w
8	1	0	0	0	→	BS	CAN	(	8	H	X	H	x
9	1	0	0	1	→	HT	EM	)	9	I	Y	I	y
A	1	0	1	0	→	LF	SUB	*	:	J	Z	J	z
B	1	0	1	1	→	VT	ESC	+	;	K	[	K	{
C	1	1	0	0	→	FF	FS	-	<	L	\	L	
D	1	1	0	1	→	CR	GS	,	=	M	]	M	}
E	1	1	1	0	→	SO	RS	.	>	N	^	N	~
F	1	1	1	1	→	SI	US	/	?	O	_	O	DEL

Control characters:

NUL = Null; DLE = Data link escape; SOH = Start Of Heading;

DC1 = Device control 1; DC2 = Device control 2; DC3 = Device control 3.

DC4 = Device control 4; STX = Start of text; ETX = End of text;

EOT = End of transmission; ENQ = Enquiry; NAK = Negative acknowledge.

ACK = Acknowledge; SYN = Synidle; BEL = Bell.

ETB = End of transmission block; BS = Backspace; CAN = Cancel.

HT = Horizontal tab; EM = End of medium; LF = Line feed; SUB = Substitute.

VT = Vertical tab; ESC = Escape; FF = Form feed; FS = File separator.

SO = Shift out; RS = Record separator; SI = Shift in; US = Unit separator.

4. Biểu diễn giá trị số trong máy tính.

4.1. Biểu diễn số nguyên.

a. Biểu diễn số nguyên không dấu:

Tất cả các số cũng như các mã ... trong máy vi tính đều được biểu diễn bằng các chữ số nhị phân. Để biểu diễn các số nguyên không dấu, người ta dùng n bit. Tương ứng với độ dài của số bit được sử dụng, ta có các khoảng giá trị xác định như sau:

Số bit	Khoảng giá trị
n bit	$0.. 2^n - 1$
8 bit	$0.. 255$ Byte
16 bit	$0.. 65535$ Word

b. Biểu diễn số nguyên có dấu:

Người ta sử dụng bit cao nhất biểu diễn dấu; bit dấu có giá trị 0 tương ứng với số nguyên dương, bit dấu có giá trị 1 biểu diễn số âm. Như vậy khoảng giá trị số được biểu diễn sẽ được tính như sau:

Số bit	Khoảng giá trị:
n bit	$2^{n-1}-1$
8 bit	$-128.. 127$ Short integer
16 bit	$-32768.. 32767$ Integer
32 bit	$-2^{31}.. 2^{31}-1$ ($-2147483648.. 2147483647$) Long integer

4.2. Biểu diễn số thực(số có dấu chấm (phẩy) động).

Có hai cách biểu diễn số thực trong một hệ nhị phân: số có dấu chấm cố định (fixed point number) và số có dấu chấm động (floating point number). Cách thứ nhất

được dùng trong những bộ VXL(micro processor) hay những bộ vi điều khiển (micro controller) cũ. Cách thứ 2 hay được dùng hiện nay có độ chính xác cao. Đối với cách biểu diễn số thực dấu chấm động có khả năng hiệu chỉnh theo giá trị của số thực. Cách biểu diễn chung cho mọi hệ đếm như sau:

$$R = m.B^e$$

Trong đó m là phần định trị, trong hệ thập phân giá trị tuyệt đối của nó phải luôn nhỏ hơn 1. Số e là phần mũ và B là cơ số của hệ đếm.

Có hai chuẩn định dạng dấu chấm động quan trọng là: chuẩn MSBIN của Microsoft và chuẩn IEEE. Cả hai chuẩn này đều dùng hệ đếm nhị phân.

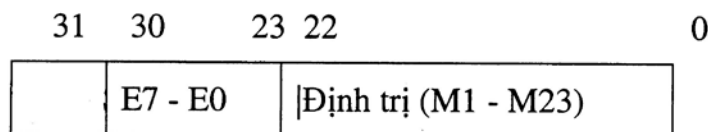
Thường dùng là theo tiêu chuẩn biểu diễn số thực của IEEE 754-1985(Institute of Electric & Electronic Engineers), là chuẩn được mọi hãng chấp nhận và được dùng trong bộ xử lý toán học của Intel. Bit dấu nằm tại vị trí cao nhất; kích thước phần mũ và khuôn dạng phần định trị thay đổi theo từng loại số thực.

Giá trị số thực IEEE được tính như sau:

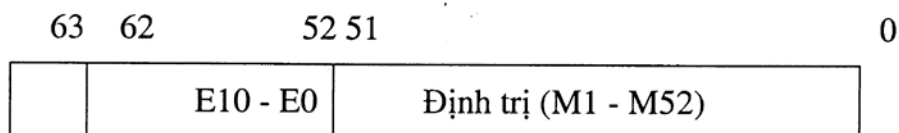
$$R = (-1)^S * (1 + M_1 * 2^{-1} + \dots + M_n * 2^{-n}) * 2^{E 7 \dots E 0 - 127}$$

Chú ý: giá trị đầu tiên M_0 luôn mặc định là 1.

- Dùng 32 bit để biểu diễn số thực, được số thực ngắn: $-3,4.10^{38} < R < 3,4.10^{38}$



- Dùng 64 bit để biểu diễn số thực, được số thực dài: $-1,7.10^{308} < R < 1,7.10^{308}$



Ví dụ tính số thực:

0100 0010 1000 1100 1110 1001 1111 1100

Phân định trị: $2^{-4} + 2^{-5} + 2^{-8} + 2^{-9} + 2^{-10} + 2^{-12} + 2^{-15} + 2^{-16} + 2^{-17} + 2^{-18} + 2^{-19} + 2^{-20} + 2^{-21} = 0,1008906$ Giá trị ngầm định là: 1,1008906 Phần mũ: $2^8 + 2^2 + 2^0 = 133$ Giá trị thực (bit cao nhất là bit dấu): $133 - 128 = 6$ Dấu: 0 = số dương	
--	--

Giá trị số thực là: $R = 1,1008906.2^6 = 70,457$.

Phương pháp đổi số thực sang số dấu phẩy động 32 bit:

- Đổi số thập phân thành số nhị phân.
- Biểu diễn số nhị phân dưới dạng $\pm 1, xxxBy$ (B: cơ số 2).
- Bit cao nhất 31: lấy giá trị 0 với số dương, 1 với số âm.
- Phần mũ y đổi sang mã excess -127 của y, được xác định bằng cách:
 $y + (7F)_{16}$.
- Phần xxx là phần định trị, được đưa vào từ bit 22..0.

Ví dụ: Biểu diễn số thực $(9,75)_{10}$ dưới dạng dấu phẩy động.

Ta đổi sang dạng nhị phân: $(9,75)_{10} = (1001.11)_2 = 1,00111B_3$.

Bit dấu: bit 31 = 0.

Mã excess - 127 của 3 là: $7F + 3 = (82)_{16} = 82H = (10000010)_2$. Được đưa vào các bit tiếp theo: từ bit 30 đến bit 23. Bit 22 luôn mặc định là 0.

Cuối cùng số thực $(9,75)_{10}$ được biểu diễn dưới dạng dấu phẩy động 32 bit như sau:

0100 0001 0001 1100 0000 0000 0000 0000

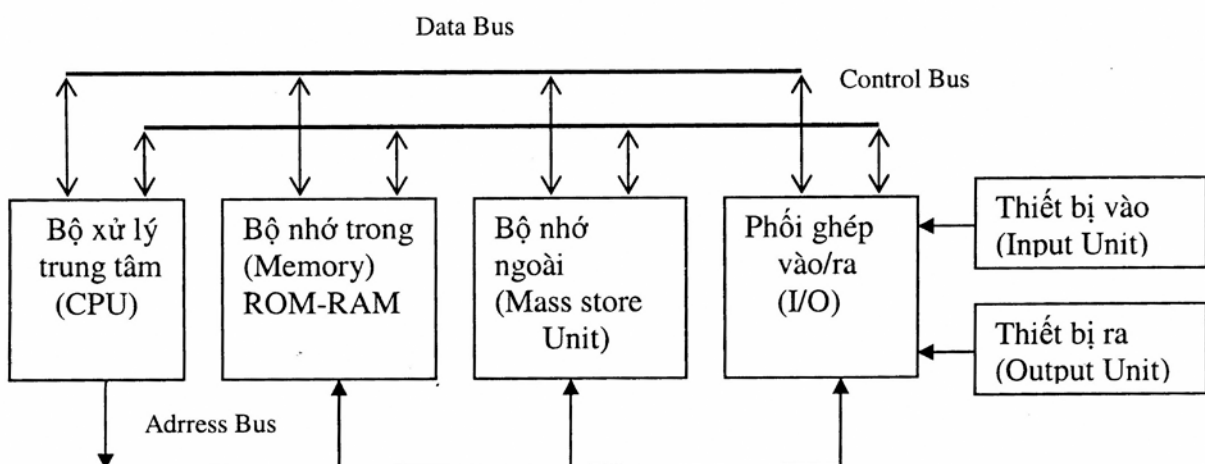
bit | 31|30 23|22 0|

§ 2. Kiến trúc một máy tính đơn giản

2.1. Giới thiệu sơ lược cấu trúc của máy vi tính.

So với từ khi ra đời, cấu trúc cơ sở của các máy vi tính ngày nay không thay đổi mấy. Mọi máy tính số đều có thể coi như được hình thành từ sáu phần chính (như hình 2-1):

Hình 2-1: Giới thiệu sơ đồ khối tổng quát của máy tính số



Trong sơ đồ này, các khối chức năng chính của máy tính số gồm:

- Khối xử lý trung tâm (central processing unit, CPU),
- Bộ nhớ trong (memory), như RAM, ROM
- Bộ nhớ ngoài, như các loại ổ đĩa, băng từ
- Khối phối ghép với các thiết bị ngoại vi (vào/ra)
- Các bộ phận đầu vào, như bàn phím, chuột, máy quét ...
- Các bộ phận đầu ra, như màn hình, máy in ...

2.2 Lịch sử phát triển của CPU

2.2.1.-BXL 4 bit

4004 là BXL đầu tiên được Intel đưa ra tháng 11 năm 1971, có tốc độ 740KHz, khả năng xử lý 0,06 triệu lệnh mỗi giây (million instructions per second - MIPS); được sản xuất trên công nghệ 10 μm , có 2.300 transistor (bóng bán dẫn), bộ nhớ mở rộng đến 640 byte.

2.2.2 BXL 8bit

8008 (năm 1972) được sử dụng trong thiết bị đầu cuối Datapoint 2200 của Computer Terminal Corporation (CTC). 8008 có tốc độ 200kHz, sản xuất trên công nghệ 10 μm , với 3.500 transistor, bộ nhớ mở rộng đến 16KB. 8080 (năm 1974) sử dụng trong máy tính Altair 8800, có tốc độ gấp 10 lần 8008 (2MHz), sản xuất trên công nghệ 6 μm , khả năng xử lý 0,64 MIPS với 6.000 transistor, có 8 bit bus dữ liệu và 16 bit bus địa chỉ, bộ nhớ mở rộng tới 64KB. 8085 có tốc độ 2MHz, sản xuất trên công nghệ 3 μm , với 6.500 transistor, có 8 bit bus dữ liệu và 16 bit bus địa chỉ, bộ nhớ mở rộng 64KB.

998

2.2.3.-BXL 16bit

80186 (năm 1982) còn gọi là IAPX 186. Sử dụng chủ yếu trong những ứng dụng nhúng, bộ điều khiển thiết bị đầu cuối. Các phiên bản của 80186 gồm 10 và 12 MHz. 80286 (năm 1982) sử dụng công nghệ 1,5 μm , 134.000 transistor, bộ nhớ mở rộng tới 16 MB. Các phiên bản của 286 gồm 6, 8, 10, 12,5, 16, 20 và 25MHz.

2.2.4. BXL 32bit vi kiến trúc NetBurst (NetBurst micro-architecture)

Intel386 gồm các họ 386DX, 386SX và 386SL. Intel386DX là BXL 32 bit đầu tiên Intel giới thiệu vào năm 1985, 386 sử dụng các thanh ghi 32 bit, có thể truyền 32 bit dữ liệu cùng lúc trên bus dữ liệu và dùng 32 bit để xác định địa chỉ. Cũng như BXL 80286, 80386 hoạt động ở 2 chế độ: real mode và protect mode.

386SL (năm1990) được thiết kế cho thiết bị di động, sử dụng công nghệ 1 μm , 855.000 transistor, bộ nhớ mở rộng 4GB; gồm các phiên bản 16, 20, 25 MHz. 486DX sử dụng công nghệ 1 μm , 1,2 triệu transistor, bộ nhớ mở rộng 4GB; gồm các phiên bản 25 MHz, 33 MHz và 50 MHz (0,8 μm). Pentium sử dụng công nghệ 0,8 μm chứa 3,1 triệu transistor, có các tốc độ 60, 66 MHz (socket 4 273 chân, PGA).

Các phiên bản 75, 90, 100, 120 MHz sử dụng công nghệ 0,6 μm chứa 3,3 triệu transistor (socket 7, PGA). Phiên bản 133, 150, 166, 200 sử dụng công nghệ 0,35 μm chứa 3,3 triệu transistor (socket 7, PGA). Pentium MMX sử dụng công nghệ 0,35 μm chứa 4,5 triệu transistor, có các tốc độ 166, 200, 233 MHz (Socket 7, PGA).

2.2.5. Pentium Pro:

Nối tiếp sự thành công của dòng Pentium, Pentium Pro được Intel giới thiệu vào tháng 9 năm 1995, sử dụng công nghệ 0,6 và 0,35 μm chứa 5,5 triệu transistor, socket 8 387 chân, Dual SPGA, hỗ trợ bộ nhớ RAM tối đa 4GB.

2.2.6. BXL Pentium II

Đầu tiên, tên mã Klamath, sản xuất trên công nghệ 0,35 μm , có 7,5 triệu transistor, bus hệ thống 66 MHz, gồm các phiên bản 233, 266, 300MHz. Pentium II, tên mã Deschutes, sử dụng công nghệ 0,25 μm , 7,5 triệu transistor, gồm các phiên bản 333MHz (bus hệ thống 66MHz), 350, 400, 450 MHz (bus hệ thống 100MHz). Celeron (năm 1998) được “rút gọn” từ kiến trúc BXL Pentium II, dành cho dòng máy cấp thấp.

2.2.7. Pentium III (năm 1999)

Bổ sung 70 lệnh mới (Streaming SIMD Extensions - SSE) giúp tăng hiệu suất hoạt động của BXL trong các tác vụ xử lý hình ảnh, audio, video và nhận dạng giọng nói. Pentium III gồm các tên mã Katmai, Coppermine và Tualatin. Coppermine sử dụng công nghệ 0,18 μm , 28,1 triệu transistor, bộ nhớ đệm L2 256 KB được tích hợp bên trong nhằm tăng tốc độ xử lý. Tualatin áp dụng công nghệ 0,13 μm có 28,1 triệu transistor, bộ nhớ đệm L1 32KB, L2 256 KB hoặc 512 KB tích hợp bên trong BXL, socket 370 FC-PGA (Flip-chip pin grid array), bus hệ thống 133 MHz. Có các tốc độ như 1133, 1200, 1266, 1333, 1400 MHz. Celeron Coppermine (năm 2000) được “rút gọn” từ kiến trúc BXL Pentium III Coppermine, còn gọi là Celeron II, được bổ sung 70 lệnh SSE. Sử dụng công nghệ 0,18 μm có 28,1 triệu transistor, bộ nhớ đệm L1 32KB, L2 256 KB tích hợp bên trong BXL, socket 370 FC-PGA, Có các tốc độ như 533, 566, 600, 633, 667, 700, 733, 766, 800 MHz (bus 66 MHz), 850, 900, 950, 1000, 1100, 1200, 1300 MHz (bus 100 MHz). Tualatin Celeron (Celeron S) (năm 2000) được “rút gọn” từ kiến trúc BXL Pentium III Tualatin, áp dụng công nghệ 0,13 μm , bộ nhớ đệm L1 32KB, L2 256 KB tích hợp, socket 370 FC-PGA, bus hệ thống 100 MHz, gồm các tốc độ 1,0, 1,1, 1,2, 1,3 và 1,4 GHz.

2.2.8. Pentium 4

Intel Pentium 4 (P4) là BXL thế hệ thứ 7 dòng x86 phổ thông, được giới thiệu vào tháng 11 năm 2000. P4 sử dụng vi kiến trúc NetBurst có thiết kế hoàn toàn mới so với các BXL cũ (PII, PIII và Celeron sử dụng vi kiến trúc P6). Một số công nghệ nổi bật được áp dụng trong vi kiến trúc NetBurst như Hyper Pipelined Technology mở rộng số hàng lệnh xử lý, Execution Trace Cache tránh tình trạng lệnh bị chậm trễ khi chuyển từ bộ nhớ đến CPU, Rapid Execution Engine tăng tốc bộ đồng xử lý toán học, bus hệ thống (system bus) 400 MHz và 533 MHz; các công nghệ

Advanced Transfer Cache, Advanced Dynamic Execution, Enhanced Floating point và Multimedia Unit, Streaming SIMD Extensions 2 (SSE2) cũng được cải tiến nhằm tạo ra những BXL tốc độ cao hơn, khả năng tính toán mạnh hơn, xử lý đa phương tiện tốt hơn. Pentium 4 đầu tiên (tên mã Willamette) xuất hiện cuối năm 2000 đặt dấu chấm hết cho "triều đại" Pentium III. Willamette sản xuất trên công nghệ 0,18 μm , có 42 triệu transistor (nhiều hơn gần 50% so với Pentium III), bus hệ thống (system bus) 400 MHz, bộ nhớ đệm tích hợp L2 256 KB, socket 423 và 478. P4 Willamette có một số tốc độ như 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2,0 GHz. P4 Northwood. Xuất hiện vào tháng 1 năm 2002, được sản xuất trên công nghệ 0,13 μm , có khoảng 55 triệu transistor, bộ nhớ đệm tích hợp L2 512 KB, socket 478. Northwood có 3 dòng gồm Northwood A (system bus 400 MHz), tốc độ 1,6, 1,8, 2,0, 2,2, 2,4, 2,5, 2,6 và 2,8 GHz. Northwood B (system bus 533 MHz), tốc độ 2,26, 2,4, 2,53, 2,66, 2,8 và 3,06 GHz (riêng 3,06 GHz có hỗ trợ công nghệ siêu phân luồng Hyper Threading - HT). Northwood C (system bus 800 MHz, tắt cả hỗ trợ HT), gồm 2,4, 2,6, 2,8, 3,0, 3,2, 3,4 GHz. P4 Prescott (năm 2004). Là BXL đầu tiên Intel sản xuất theo công nghệ 90 nm, kích thước vi mạch giảm 50% so với P4 Willamette. Điều này cho phép tích hợp nhiều transistor hơn trên cùng kích thước (125 triệu transistor so với 55 triệu transistor của P4 Northwood), tốc độ chuyển đổi của transistor nhanh hơn, tăng khả năng xử lý, tính toán. Dung lượng bộ nhớ đệm tích hợp L2 của P4 Prescott gấp đôi so với P4 Northwood (1MB so với 512 KB). Ngoài tập lệnh MMX, SSE, SSE2, Prescott được bổ sung tập lệnh SSE3 giúp các ứng dụng xử lý video và game chạy nhanh hơn. Đây là giai đoạn "giao thời" giữa socket 478 - 775LGA, system bus 533 MHz - 800 MHz và mỗi sản phẩm được đặt tên riêng khiến người dùng càng bối rối khi chọn mua. Prescott A (FSB 533 MHz) có các tốc độ 2,26, 2,4, 2,66, 2,8 (socket 478), Prescott 505 (2,66 GHz), 505J (2,66 GHz), 506 (2,66 GHz), 511 (2,8 GHz), 515 (2,93 GHz), 515J (2,93 GHz), 516 (2,93 GHz), 519J (3,06 GHz), 519K (3,06 GHz) sử dụng socket 775LGA. Prescott E, F (năm 2004) có bộ nhớ đệm L2 1 MB (các phiên bản sau được mở rộng 2 MB), bus hệ thống 800 MHz. Ngoài tập lệnh MMX, SSE, SSE2, SSE3 tích hợp, Prescott E, F còn hỗ trợ công nghệ siêu phân luồng, một số phiên bản sau có hỗ trợ tính toán 64 bit.

Dòng sử dụng socket 478 gồm Pentium 4 HT 2.8E (2,8 GHz), 3.0E (3,0 GHz), 3.2E (3,2 GHz), 3.4E (3,4 GHz). Dòng sử dụng socket 775LGA gồm Pentium 4 HT 3.2F, 3.4F, 3.6F, 3.8F với các tốc độ tương ứng từ 3,2 GHz đến 3,8 GHz, Pentium 4 HT 517, 520, 520J, 521, 524, 530, 530J, 531, 540, 540J, 541, 550, 550J, 551, 560, 560J, 561, 570J, 571 với các tốc độ từ 2,8 GHz đến 3,8 GHz.

2.2.9. BXL Celeron

BXL Celeron được thiết kế với mục tiêu dung hòa giữa công nghệ và giá cả, đáp ứng các yêu cầu phổ thông như truy cập Internet, Email, chat, xử lý các ứng dụng văn phòng. Celeron Willamette 128 (2002), bản "rút gọn" từ P4 Willamette, sản xuất trên công nghệ 0,18 μm , bộ nhớ đệm L2 128 KB, bus hệ thống 400 MHz, socket 478. Celeron Willamette 128 hỗ trợ tập lệnh MMX, SSE, SSE2. Một số BXL thuộc dòng này như Celeron 1.7 (1,7 GHz) và Celeron 1.8 (1,8 GHz). Celeron NorthWood 128,

"rút gọn" từ P4 Northwood, công nghệ 0,13 μm , bộ nhớ đệm tích hợp L2 128 KB, bus hệ thống 400 MHz, socket 478. Celeron NorthWood 128 cũng hỗ trợ các tập lệnh MMX, SSE, SSE2, gồm Celeron 1.8A, 2.0, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.8 tương ứng với các tốc độ từ 1,8 GHz đến 2,8 GHz. Celeron D (Presscott 256), được xây dựng từ nền tảng P4 Prescott, sản xuất trên công nghệ 90nm, bộ nhớ đệm tích hợp L2 256 KB (gấp đôi dòng Celeron NorthWood), bus hệ thống 533 MHz, socket 478 và 775LGA. Ngoài các tập lệnh MMX, SSE, SSE2, Celeron D hỗ trợ tập lệnh SSE3, một số phiên bản sau có hỗ trợ tính toán 64 bit. Celeron D gồm 310, 315, 320, 325, 325J, 326, 330, 330J, 331, 335, 335J, 336, 340, 340J, 341, 345, 345J, 346, 350, 351, 355 với các tốc độ tương ứng từ 2,13 GHz đến 3,33 GHz.

2.2.10. Pentium 4 Extreme Edition

Pentium 4 Extreme Edition (P4EE) xuất hiện vào tháng 9 năm 2003, là BXL được Intel "ưu ái" dành cho game thủ và người dùng cao cấp. P4EE được xây dựng từ BXL Xeon dành cho máy chủ và trạm làm việc. Ngoài công nghệ HT "đỉnh đám" thời bấy giờ, điểm nổi bật của P4EE là bổ sung bộ nhớ đệm L3 2 MB. Phiên bản đầu tiên của P4 EE (nhân Gallatin) sản xuất trên công nghệ 0,13 μm , bộ nhớ đệm L2 512 KB, L3 2 MB, bus hệ thống 800 MHz, sử dụng socket 478 và 775LGA, gồm P4 EE 3.2 (3,2 GHz), P4 EE 3.4 (3,4 GHz).

2.2.11.BXL 64 bit, vi kiến trúc NETBURST

P4 Prescott (năm 2004) Vi kiến trúc NetBurst 64 bit (Extended Memory 64 Technology - EM64T) đầu tiên được Intel sử dụng trong BXL P4 Prescott (tên mã Prescott 2M). Prescott 2M cũng sử dụng công nghệ 90 nm, bộ nhớ đệm L2 2 MB, bus hệ thống 800 MHz, socket 775LGA. Ngoài các tập lệnh MX, SSE, SSE2, SSE3, công nghệ HT và khả năng tính toán 64 bit, Prescott 2M (trừ BXL 620) có hỗ trợ công nghệ Enhanced SpeedStep để tối ưu tốc độ làm việc nhằm tiết kiệm điện năng. Các BXL 6x2 có thêm công nghệ ảo hóa (Virtualization Technology). Prescott 2M có một số tốc độ như P4 HT 620 (2,8 GHz), 630 (3,0 GHz), 640 (3,2 GHz), 650 (3,4 GHz), 660, 662 (3,6 GHz) và 670, 672 (3,8 GHz).

2.2.12. Pentium D (năm 2005)

Pentium D (tên mã Smithfield, 8xx) là BXL lõi kép (dual core) đầu tiên của Intel, được cải tiến từ P4 Prescott nên cũng gặp một số hạn chế như hiện tượng thất cổ chai do băng thông BXL ở mức 800 MHz (400 MHz cho mỗi lõi), Cùng sử dụng vi kiến trúc NetBurst, Pentium D (mã Presler, 9xx) được Intel thiết kế mới trên công nghệ 65nm, 376 triệu transistor, bộ nhớ đệm L2 4 MB (2x2 MB), hiệu năng cao hơn, nhiều tính năng mới và ít tốn điện năng hơn Smithfield. Pentium D 915 và 920 tốc độ 2,8 GHz, 925 và 930 (3,0GHz), 935 và 940 (3,2 GHz), 945 và 950 (3,4 GHz), 960 (3,6GHz). Presler dòng 9x0 có hỗ trợ Virtualization Technology.

2.2.13. Pentium Extreme Edition (năm 2005)

BXL lõi kép dành cho game thủ và người dùng cao cấp. Pentium EE sử dụng nhân Smithfield, Presler của Pentium D trong đó Smithfield sử dụng công nghệ 90nm, bộ

nhớ đệm L2 được mở rộng đến 2 MB (2x1 MB), hỗ trợ tập lệnh MMX, SSE, SSE2, SSE3, công nghệ HT, Enhanced Intel SpeedStep Technology (EIST) và EM64T. Pentium 840 EE (3,20 GHz, bus hệ thống 800 MHz, socket 775LGA) là một trong những BXL thuộc dòng này.

2.2.14. BXL 64bit, kiến trúc Core

Tại diễn đàn IDF đầu năm 2006, Intel đã giới thiệu kiến trúc Intel Core với năm cải tiến quan trọng là khả năng mở rộng thực thi động (Wide Dynamic Execution), tính năng quản lý điện năng thông minh (Intelligent Power Capability), chia sẻ bộ nhớ đệm linh hoạt (Advanced Smart Cache), truy xuất bộ nhớ thông minh (Smart Memory Access) và tăng tốc phương tiện số tiên tiến (Advanced Digital Media Boost).

2.2.15. Intel Core 2 Duo

BXL lõi kép sản xuất trên công nghệ 65 nm, hỗ trợ SIMD instructions, công nghệ Virtualization Technology cho phép chạy cùng lúc nhiều HĐH, tăng cường bảo vệ hệ thống trước sự tấn công của virus (Execute Disable Bit), tối ưu tốc độ BXL nhằm tiết kiệm điện năng (Enhanced Intel SpeedStep Technology), quản lý máy tính từ xa (Intel Active Management Technology). Ngoài ra, còn hỗ trợ các tập lệnh MMX, SSE, SSE2, SSE3, SSSE3.

Core 2 Duo (tên mã Conroe) có 291 triệu transistor, bộ nhớ đệm L2 4 MB, bus hệ thống 1066 MHz, socket 775LGA. Một số BXL thuộc dòng này: E6600 (2,4 GHz), E6700 (2,66 GHz). Core 2 Duo (tên mã Allendale) E6300 (1,86 GHz), E6400 (2,13 GHz) có 167 triệu transistor, bộ nhớ đệm L2 2MB, bus hệ thống 1066 MHz, socket 775LGA. E4300 (1,8 GHz) xuất hiện năm 2007 có bộ nhớ đệm L2 2 MB, bus 800 MHz, không hỗ trợ Virtualization Technology.

2.2.16. Core 2 Extreme

BXL lõi kép dành cho game thủ sử dụng kiến trúc Core, có nhiều đặc điểm giống với BXL Core 2 như công nghệ sản xuất 65 nm, hỗ trợ các công nghệ mới Enhanced Intel SpeedStep Technology, Intel x86-64, Execute Disable Bit, Intel Active Management, Virtualization Technology, Intel Trusted Execution Technology... các tập lệnh MMX, SSE, SSE2, SSE3, SSSE3.

2.2.17. Core 2 Extreme

Core 2 Extreme (tên mã Conroe XE) (tháng 7 năm 2006) với đại diện X6800 2,93 Ghz, bộ nhớ đệm L2 đến 4 MB, bus hệ thống 1066 MHz, socket 775LGA. Cuối năm 2006, con đường phía trước của BXL tiếp tục rộng mở khi Intel giới thiệu BXL 4 nhân (Quad Core) như Core 2 Extreme QX6700, Core 2 Quad Q6300, Q6400, Q6600 và BXL 8 nhân trong vài năm tới. Chắc chắn những BXL này sẽ thỏa mãn nhu cầu người dùng đam mê công nghệ và tốc độ.

Hiện đã có loại **CPU Quad-Core** (4 nhân). Hãng AMD đã cho ra công nghệ gồm 2 bộ xử lý, mỗi bộ 2-4 nhân. Tuy nhiên loại CPU này vẫn chưa có mặt trên thị trường.

2.3 Chất liệu và công nghệ chế tạo CPU

2.3.1. Chất liệu

Gồm và organic (hữu cơ) từ dòng Thoroughbred trở đi đều làm bằng organic. Hiện tại, công nghệ được áp dụng cho các CPU Chất liệu chủ yếu chế tạo cpu AMD là ceramic à MOS (Metal Oxide Semi-Conductor - bán dẫn ôxít kim loại), dựa vào một lớp ôxít kim loại nằm trên tấm silicon kết nối bởi các đường hợp chất dẫn điện. Người ta đã cải tiến MOS thành CMOS (Complimentary MOS - MOS bổ trợ) hoạt động ở điện thế thấp. Đây là 2 công nghệ có mặt trong hầu hết các thiết bị máy tính. Để đáp ứng nhu cầu làm cho CPU ngày càng nhanh hơn, ít tiêu hao năng lượng hơn các công nghệ 0,25 -> 0,18 -> 0,13 micron lần lượt ra đời. Nhưng chính sự thu nhỏ các cầu nối trong CPU này khiến việc áp dụng MOS và CMOS trở nên ngày càng khó khăn hơn, do các cầu nối này nằm quá sát nhau nên dễ dẫn đến hiện tượng đóng điện chéo lên các cầu bên cạnh. Một nhược điểm quan trọng khác của công nghệ MOS là phần silicon ở giữa các cầu nối (có vai trò như một tụ điện) phải nạp được điện dung tối đa để có thể đóng - và lại phải thoát hết điện dung để có thể mở. Việc này tốn thời gian xử lý, và lãng phí thời gian xử lý trên CPU. Các nhà sản xuất CPU đã cải tiến MOS hiện có như việc thay oxit nhôm bằng oxit đồng làm tăng xung nhịp lên đáng kể. Nhưng để CPU có thể đạt tới tốc độ 5-10 GHz phải có một giải pháp khắc phục triệt để hơn nữa 2 nhược điểm nêu trên. Đó chính là công nghệ SOI (Silicon On Insulator). IBM đã phát triển công nghệ này từ năm 1990 cho CPU của IBM, với mục đích giảm điện năng sử dụng, tăng xung nhịp v.v...nhưng công nghệ này vẫn chưa thực sự được ứng dụng ngay cho đến cuối thế kỉ 20, khi việc tăng xung nhịp cho các dòng CPU hiện đại cần thêm các phương pháp sản xuất khác. Cải tiến SOI là điện dung của tụ silicon giữa các cầu được cực tiểu hoá làm giảm thời gian cần thiết để thoát/nạp, để mở và đóng cầu nối. Điều này giúp tăng xung nhịp lên rất nhiều. Sở dĩ SOI làm được điều đó là nhờ việc chèn vào giữa tấm silicon một lớp vật liệu cách điện và để lại một phần silicon nhỏ ở giữa các cầu nối. Lớp vật liệu cách điện này là một dạng của ôxít silicon được tạo ra bằng kĩ thuật SIMOX (Seperation by Implantation of Oxygen) - khí ôxi được ép lên bề mặt của silicon wafer ở áp suất và nhiệt độ cao, khi đó silicon phản ứng với ôxi tạo nên 1 lớp ôxít silicon bám vào silicon wafer bên dưới.

SOI sẽ không thay thế hoàn toàn MOS/CMOS mà chỉ tối ưu hoá cho hai công nghệ này:

- CPU dùng SOI sẽ nhanh hơn đến 30% so với CPU dùng MOS/CMOS nếu có cùng một xung đồng hồ như nhau.
- Yêu cầu về điện năng thấp hơn nhiều so với MOS/CMOS (ít hơn khoảng 50%), CPU sẽ chạy mát hơn - vượt qua một trở ngại lớn của việc nâng tốc độ các bộ xử lý.
- Cho phép thu nhỏ công nghệ sản xuất CPU xuống 0.09 micron hay thấp hơn cùng với SOI có nghĩa rằng các bộ vi xử lý sẽ được tăng tốc rất nhanh và tốc độ 5-10GHz sẽ sớm đạt được. Thế nhưng SOI cần có silicon đạt độ nguyên chất 100% - thứ mà công nghệ hiện nay chưa sản xuất được. Isonics là 1 công ty đang nghiên cứu sản xuất loại

silicon wafer này. AMD thực sự trông đợi vào SOI để khắc phục những nhược điểm của CPU như tiêu tốn nhiều điện năng và chạy nóng hơn. bộ xử lý K8 của IBM, hay còn gọi là Hammer dùng công nghệ SOI đang được mong đợi. Nội lực công nghệ - HyperTransport, Cool'n'Quiet. AMD đặc biệt ưu ái CPU 64 bit với công nghệ 'siêu chuyển' HyperTransport và tự điều chỉnh hoạt động Cool'n'Quiet. HyperTransport giúp việc truyền thông tin giữa các chip (cầu nam, cầu bắc, BXL, bộ nhớ,...) nhanh hơn, khả năng 'nói chuyện' với một chip hoặc thiết bị khác nhanh hơn với lượng tiêu thụ lớn hơn. HyperTransport làm cho đường truyền rộng hơn, do đó tốc độ truyền nhanh và nhiều hơn. Công nghệ này có thể áp dụng cho tất cả băng thông của bo mạch chủ, từ chipset đến BXL, bộ nhớ, AGP, PCI,...Cool'n'Quiet là một cải tiến khác dành cho dòng BXL 64 bit, tốc độ và điện năng tiêu thụ của BXL sẽ được điều chỉnh tự động. Nếu có ít ứng dụng được chạy (BXL xử lý ít) thì Cool'n'Quiet sẽ giảm tốc độ và điện thế BXL, ngược lại, khi cần xử lý nhiều thì BXL sẽ được tăng tốc độ và điện thế.

2.4 Nguyên tắc hoạt động của CPU

CPU (Central Processing Unit) – cũng được gọi là microprocessor hay processor – là một đơn vị xử lý dữ liệu trung tâm. Cách nó xử lý dữ liệu như thế nào hoàn toàn phụ thuộc vào chương trình được viết từ trước. Chương trình nói chung có thể là một bảng tính, một bộ xử lý từ hay một game nào đó. Nó chỉ tuân theo các thứ tự (được gọi là các chỉ lệnh hay các lệnh) có bên trong chương trình.

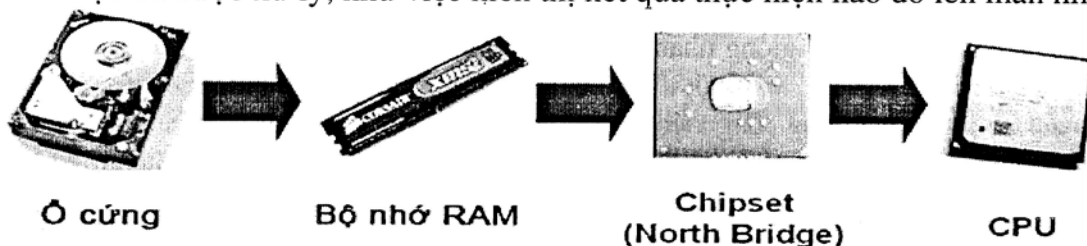
Khi một chương trình nào đó được chạy thì thứ tự được thực hiện như sau:

a. Chương trình đã lưu bên trong ổ đĩa cứng sẽ được đưa vào bộ nhớ RAM. Ở đây chương trình chính là một loạt các chỉ lệnh đối với CPU.

b. CPU sử dụng mạch phần cứng được gọi là memory controller để tải dữ liệu chương trình từ bộ nhớ RAM.

c. Lúc đó dữ liệu bên trong CPU sẽ được xử lý.

d. Những gì diễn ra tiếp theo sẽ phụ thuộc vào chương trình vừa được nạp. CPU có thể tiếp tục tải và thực thi chương trình hoặc có thể thực hiện một công việc nào đó với dữ liệu đã được xử lý, như việc hiển thị kết quả thực hiện nào đó lên màn hình.



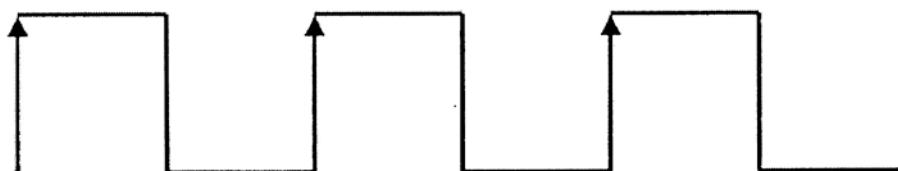
Hình 2.4: Dữ liệu lưu được đưa vào CPU

Sự truyền tải dữ liệu giữa ổ đĩa cứng và bộ nhớ RAM được thực hiện mà không sử dụng đến CPU, như vậy nó sẽ làm cho hệ thống hoạt động nhanh hơn. Phương pháp này được gọi là bus mastering hay DMA (Direct Memory Access). Các bộ vi xử lý của AMD dựa trên sockets 754, 939 và 940 (Athlon 64, Athlon 64 X2, Athlon 64 FX, Opteron và

một số mô hình Sempron) có một memory controller được nhúng bên trong. Điều đó có nghĩa rằng với các bộ vi xử lý này, CPU truy cập trực tiếp bộ nhớ RAM.

2.4.1.Clock

Clock chính là một tín hiệu được sử dụng để đồng bộ hóa mọi thứ bên trong máy tính. Hãy xem trong hình 2.4.1, đây chính là một xung clock điển hình: nó là một xung hình vuông biến thiên ở mức “0” và “1” với một tốc độ được cố định. Trên hình vẽ ta có thể thấy 3 chu kỳ của xung clock này. Bắt đầu của mỗi một chu kỳ khi tín hiệu clock biến thiên từ “0” lên “1”; chúng được đánh dấu nó bằng một mũi tên. Tín hiệu clock được đo theo đơn vị có tên gọi là Hertz (Hz), đây là số chu kỳ clock trong mỗi giây đồng hồ. Một xung clock 100MHz có nghĩa là trong một giây đồng hồ có 100 triệu chu kỳ xung nhịp.



Hình 2.4.1: Tín hiệu xung clock

Trong máy tính, tất cả các bộ định thời đều được đo dưới dạng các chu kỳ clock. Ví dụ, một bộ nhớ RAM có độ trễ là “5” thì điều đó có nghĩa là nó sẽ giữ chậm 5 chu kỳ xung nhịp để thực hiện công việc cung cấp dữ liệu. Trong CPU, tất cả các chỉ lệnh giữ chậm một số chu kỳ xung clock nào đó để được thực thi. Ví dụ, một chỉ lệnh nào đó có thể được giữ chậm đến 7 chu kỳ xung clock để được thực thi xong.

CPU biết được bao nhiêu chu kỳ xung clock mà mỗi chỉ lệnh cần, nó biết được điều này bởi CPU giữ một bảng liệt kê các thông tin này. Chính vì vậy nếu CPU có hai chỉ lệnh được thực thi và nó biết rằng chỉ lệnh đầu tiên sẽ giữ chậm 7 chu kỳ xung clock để thực thi thì nó sẽ tự động thực thi chỉ lệnh kế tiếp vào chu kỳ clock thứ 8. Rõ ràng đây là một cách lý giải chung cho CPU với một khối thực thi – các bộ vi xử lý hiện đại có một số khối thực thi làm việc song song và nó có thể thực thi chỉ lệnh thứ hai tại cùng thời điểm với chỉ lệnh đầu. Điều này được gọi là kiến trúc “superscalar”.

Nếu so sánh hai CPU giống nhau, CPU nào chạy ở tốc độ clock cao hơn sẽ nhanh hơn. Trong trường hợp này, với một tốc độ clock cao hơn, thời gian giữa mỗi chu kỳ clock sẽ ngắn hơn, vì vậy những công việc sẽ được thực thi tốn ít thời gian hơn và hiệu suất sẽ cao hơn. Tuy nhiên khi so sánh hai bộ vi xử lý khác nhau thì điều này hoàn toàn không đúng.

Nếu ta lấy hai bộ vi xử lý có kiến trúc khác nhau – ví dụ, khác nhau về nhà sản xuất như Intel và AMD – những thứ bên trong hai CPU này là hoàn toàn khác nhau. Như đã đề cập, mỗi chỉ lệnh cần đến một số chu kỳ clock nhất định để được thực thi. Chúng ta hãy nói rằng bộ vi xử lý “A” cần đến 7 chu kỳ clock để thực thi một chỉ lệnh nào đó và bộ vi xử lý “B” cần 5 chu kỳ clock để thực hiện một chỉ lệnh tương tự. Nếu chúng đang chạy với cùng một tốc độ clock thì bộ vi xử lý “B” sẽ nhanh hơn, vì nó có thể xử lý chỉ lệnh này tốn ít thời gian hơn. Với các CPU hiện đại, có nhiều vấn đề cần phải

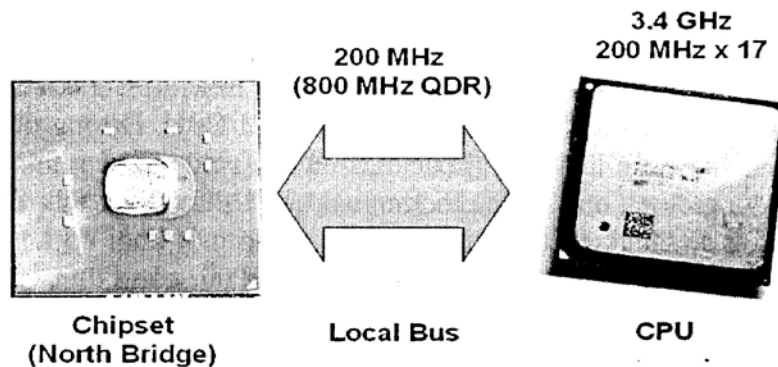
xem xét đến hiệu suất này, vì các CPU có số lượng khối thực thi khác nhau, kích thước cache khác nhau, các cách truyền tải dữ liệu bên trong CPU cũng khác nhau, cách xử lý các chỉ lệnh bên trong các khối thực thi và tốc độ clock khác nhau với thế giới thực bên ngoài,...

Khi tần hiệu clock của bộ vi xử lý cao thì có một vấn đề mà chúng ta gặp phải. Bo mạch chủ, nơi mà bộ vi xử lý được cài đặt không thể làm việc bằng cách sử dụng cùng tần hiệu clock. Nếu xem bo mạch chủ, ta sẽ thấy một số đường và rãnh. Các đường và rãnh này là những mạch in nối một số mạch của máy tính. Vấn đề ở đây là với tốc độ clock cao, các dây mạch in này sẽ bắt đầu làm việc như anten, chính vì vậy tín hiệu, thay vì đến vị trí cần đến ở phía cuối đầu dây lại biến mất. được truyền đi như các sóng vô tuyến.

2.4.2 External Clock

Vì vậy các nhà sản xuất CPU đã bắt đầu sử dụng một khái niệm mới, khái niệm được gọi là nhân xung clock, ứng dụng này bắt đầu được sử dụng trong bộ vi xử lý 486DX2. Với cơ chế này (được sử dụng trong tất cả các CPU ngày nay), CPU có một clock ngoài (external clock) được sử dụng khi truyền tải dữ liệu vào ra bộ nhớ RAM (sử dụng north bridge chip) và một clock trong cao hơn.

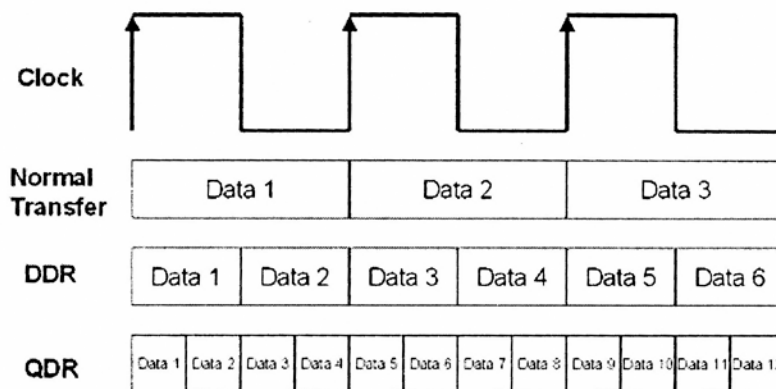
Để đưa ra một ví dụ thực, trong số 3.4 GHz Pentium 4 thì con số “3.4 GHz” chính là clock trong của CPU, clock này đạt được bằng cách nhân 17 với clock ngoài là 200 của nó. Mô phỏng ví dụ này trong hình 2.4.2



Hình 2.4.2: Clock trong và ngoài trên Pentium 4 3.4 GHz.

Sự khác nhau lớn giữa clock trong và clock ngoài trên các CPU hiện đại là cách vượt qua nhược điểm từ tính như đã nói trên để tăng hiệu suất máy tính. Tiếp tục với ví dụ về Pentium 4 3.4 GHz ở trên, nó phải giảm tốc độ của nó đi 17 lần khi thực hiện đọc dữ liệu từ bộ nhớ RAM! Trong suốt quá trình này, nó làm việc như một CPU với tốc độ 200MHz.

Một số kỹ thuật được sử dụng để tối thiểu hóa ảnh hưởng của sự khác nhau clock này. Một trong số chúng là sử dụng cache nhớ bên trong CPU. Phương pháp khác là truyền tải nhiều khối dữ liệu trên mỗi một chu kỳ clock. Các bộ vi xử lý của hai hãng Intel và AMD đều sử dụng tính năng này, tuy nhiên trong khi CPU của AMD truyền tải hai dữ liệu trên một chu kỳ clock thì các CPU của Intel truyền tải 4 dữ liệu trên mỗi chu kỳ.



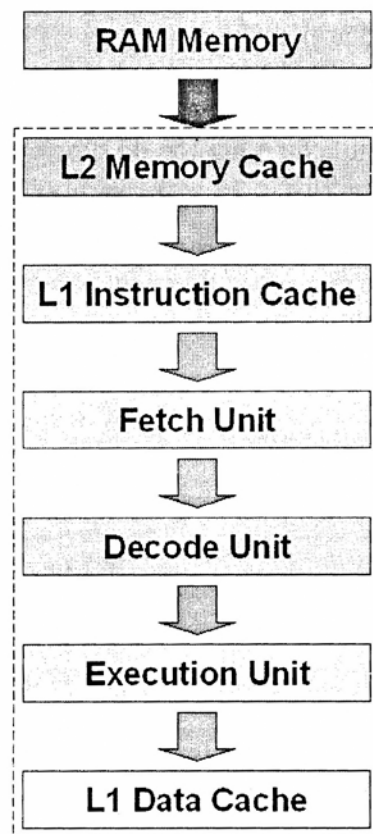
Hình 2.4.3: Truyền tải nhiều dữ liệu trên mỗi chu kỳ clock

Chính vì điều đó nên các CPU của AMD được liệt vào loại có tốc độ gấp hai clock ngoài thực. Ví dụ, một CPU của AMD với external clock là 200MHz được liệt vào CPU có clock ngoài là 400MHz. Điều tương tự cũng được áp dụng đối với các CPU của Intel, với external clock là 200MHz thì CPU của nó sẽ có tốc độ clock ngoài là 800MHz.

Kỹ thuật truyền tải hai dữ liệu trên mỗi một chu kỳ clock được gọi là DDR (Dual Data Rate), còn kỹ thuật truyền tải 4 dữ liệu trên một chu kỳ clock được gọi là QDR (Quad Data Rate).

2.4.3 Sơ đồ khối của một CPU

Trên hình 2.4.4 sơ đồ khối cơ bản của một CPU hiện đại. Có nhiều sự khác nhau giữa các kiến trúc của AMD và Intel. Việc hiểu được các kiến thức cơ bản này sẽ là một bước để ta có thể hiểu được cách các CPU của Intel và AMD làm việc như thế nào và sự khác nhau giữa chúng. Dòng nét chấm trên hình 2.4.4 thể hiện phần “body” của CPU, vì bộ nhớ RAM được đặt bên ngoài CPU. Đường dữ liệu giữa bộ nhớ RAM và CPU thường là 64-bit (hoặc 128-bit khi sử dụng cấu hình bộ nhớ kênh dual), đang sử dụng clock nhớ hoặc clock ngoài của CPU (clock thấp). Số lượng bit đã sử dụng và tốc độ clock có thể được kết hợp trong một khối có tên gọi là tốc độ truyền tải, tính theo MB/s. Để tính toán tốc độ truyền tải, công thức được thực hiện tính tốc độ này bằng số bit x clock/8. Với hệ thống sử dụng các bộ nhớ DDR400 trong cấu hình kênh đơn (64 bit) thì tốc độ truyền tải sẽ là 3.200MB/s, còn với hệ thống tương tự sử dụng các bộ nhớ kênh dual (128 bit) sẽ có tốc độ truyền tải bộ nhớ là 6.400 MB/s.



Hình 2.4.4. Sơ đồ khối cơ bản của một CPU

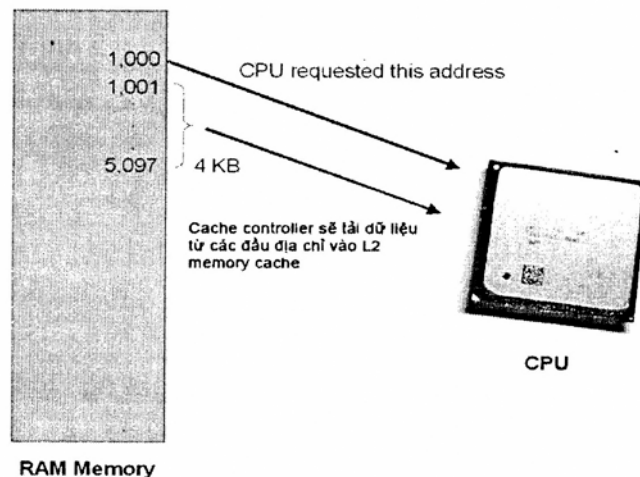
Tất cả các mạch bên trong phần đánh dấu chấm chạy ở tốc độ clock trong của CPU. Phụ thuộc vào CPU mà một số phần bên trong có nó có thể chạy ở tốc độ clock cao hơn. Cũng vậy, đường dữ liệu giữa các khối CPU có thể rộng hơn, nghĩa là truyền tải nhiều bit hơn trên mỗi chu kỳ clock 64 và 128. Ví dụ, đường dữ liệu giữa bộ nhớ cache L2 và cache chỉ lệnh L1 trên các bộ vi xử lý hiện đại thường là 256 bit. Số bit được truyền tải trên mỗi chu kỳ clock càng cao thì sự truyền tải sẽ được thực hiện càng nhanh (hay nói cách khác, tốc độ truyền tải sẽ cao hơn). Trên hình 2.4.4, mũi tên giữa bộ nhớ RAM và cache nhớ L2; mũi tên giữa các khối khác để diễn tả tốc độ clock khác nhau và bề rộng của đường dữ liệu đã sử dụng.

2.4.4 Memory Cache

Memory Cache là một kiểu bộ nhớ hiệu suất cao, cũng được gọi là bộ nhớ tĩnh. Kiểu bộ nhớ đã sử dụng trên bộ nhớ RAM chính của máy tính được gọi là bộ nhớ động. Bộ nhớ tĩnh tiêu tốn nhiều năng lượng điện hơn, đắt hơn và có kích thước vật lý lớn hơn so với bộ nhớ động, tuy nhiên nó lại chạy nhanh hơn. Nó có thể làm việc với cùng tốc độ clock của CPU, điều mà bộ nhớ động không thể thực hiện được. Khi CPU cần nạp dữ liệu ở ngoài, nó phải làm việc ở tốc độ clock thấp hơn do vậy mà kỹ thuật cache nhớ được sử dụng ở đây để khắc phục nhược điểm này. Khi CPU nạp dữ liệu từ một vị trí nhớ nào đó thì mạch điều khiển memory cache controller nạp vào cache nhớ một khối dữ liệu bên dưới vị trí hiện hành mà CPU đã nạp. Vì các chương trình được thực hiện theo thứ tự nên vị trí nhớ tiếp theo mà CPU sẽ yêu cầu có thể là bị trí ngay dưới vị trí nhớ mà nó đã nạp. Do memory cache controller đã nạp rất

nhiều dữ liệu dưới vị trí nhớ đầu tiên được đọc bởi CPU nên dữ liệu kế tiếp sẽ ở bên trong cache nhớ, chính vì vậy CPU không cần phải thực hiện thao tác lấy dữ liệu bên ngoài: nó đã được nạp vào bên trong cache nhớ nhúng trong CPU, chính vì những trong CPU mà chúng có thể truy cập bằng tốc độ clock trong.

Cache controller luôn luôn quan sát các vị trí nhớ đã và đang được nạp dữ liệu từ một vài vị trí nhớ sau khi vị trí nhớ vừa được đọc. Một ví dụ thực tế, nếu một CPU đã nạp dữ liệu được lưu tại địa chỉ 1.000 thì cache controller sẽ nạp dữ liệu từ “n” địa chỉ sau địa chỉ 1.000. Số “n” được gọi là trang; nếu một bộ vi xử lý này làm việc với



4KB

Hình 2.4.5: Memory cache controller làm việc như thế nào

trang (giá trị điển hình) thì nó sẽ nạp dữ liệu từ các địa chỉ 4.096 dưới vị trí nhớ hiện hành đang được nạp (địa chỉ 1.000 trong ví dụ). 1KB bằng 1.024 byte, do đó là 4,096 chứ không phải 4,000. Chúng tôi đã thể hiện ví dụ này trên hình 2.4.5. Memory cache càng lớn thì cơ hội cho dữ liệu yêu cầu bởi CPU ở đây càng cao, chính vì vậy CPU sẽ giảm sự truy cập trực tiếp vào bộ nhớ RAM, do đó hiệu suất hệ thống tăng (hãy nên nhớ rằng khi CPU cần truy cập trực tiếp vào bộ nhớ RAM thì nó phải thực hiện ở tốc độ clock thấp hơn nên giảm hiệu suất của toàn hệ thống).

Chúng ta gọi là “hit” khi CPU nạp một dữ liệu yêu cầu từ cache và “miss” nếu dữ liệu yêu cầu không có ở đó và CPU cần phải truy cập vào bộ nhớ RAM của hệ thống. L1 và L2 tương ứng là “Level 1” và “Level 2”, được đại diện cho khoảng cách chúng cách lõi CPU (khối thực thi). Một sự ngờ vực hay có ở đây là tại sao có đến 3 bộ nhớ Cache (L1 data cache, L1 instruction cache và L2 cache). Hãy chú ý trên hình 2.4.5 và thấy được rằng L1 instruction cache làm việc như một “input cache”, trong khi đó L1 data cache làm việc như một “output cache”. L1 instruction cache – thường nhỏ hơn L2 cache – chỉ hiệu quả khi chương trình bắt đầu lặp lại một phần nhỏ của nó (loop), vì các chỉ lệnh yêu cầu sẽ gần hơn với khối tìm nạp. Trên trang chi tiết kỹ thuật của một CPU, L1 cache có thể được thể hiện bằng một hình ảnh hoàn toàn khác. Một số nhà máy sản xuất liệt kê hai L1 cache riêng biệt (đôi khi gọi cache chỉ lệnh là “I” và cache dữ liệu là “D”), một số hãng ghi số lượng của cả hai là 128 KB nhưng điều đó có nghĩa là 64 KB cho cache chỉ lệnh và 64 KB cho cache dữ liệu.