

INTRODUÇÃO À ARQUITETURA DE COMPUTADORES

3.0 Índice

3.1	Introdução	2
3.2	Componentes básicos de um computador	2
3.3	Processador	2
3.3.1	Unidade Lógica e Aritmética	4
3.3.2	Unidade de Controle (UC)	7
3.3.3	Registradores	7
3.3.4	Clock	7
3.4	Processadores CISC e RISC	8
3.5	Memória	8
3.5.1	Memória Principal	9
3.5.2	Tecnologias de memória RAM	11
3.5.3	Memória Secundária	15
3.5.4	Memória Cache	19
3.6	Barramentos	21
3.6.1	Barramento do processador	21
3.6.2	Barramento de Cache	22
3.6.3	Barramento da Memória	22
3.6.4	Barramento de E/S	23
3.6.5	Barramento ISA	23
3.6.6	Barramento MCA	24
3.6.7	Barramento EISA	24
3.6.8	Local bus	24
3.6.9	VESA Local Bus	26
3.6.10	Barramento PCI	26
3.6.11	Barramento PC-Card (PCMCIA)	27
3.6.12	USB	28
3.6.13	Firewire (IEEE 1394)	28
3.7	Dispositivos de Entrada/Saída	28

3.1 Introdução

Como foi mencionado no capítulo 1, tem-se observado uma atuação cada vez maior dos computadores nas diversas atividades do nosso dia a dia. As operações bancárias, as telecomunicações e o manuseio de muitos aparelhos eletrodomésticos são exemplos claros das facilidades trazidas pela utilização dos computadores, isto sem falar em aplicações mais clássicas, como os sistemas de reservas de passagens aéreas e a previsão meteorológica.

A evolução da informática foi caracterizada pelo desenvolvimento de computadores com características as mais diversas, traduzidas pelos diferentes parâmetros, cada vez mais conhecidos da maioria de usuários de computador: a CPU adotada, a capacidade de memória, a capacidade do disco rígido, a existência de memória cache e outros menos conhecidos. A definição destes parâmetros e a forma como os diversos componentes de um computador são organizados, define aquilo que é conhecido por *arquitetura de computador* e vai determinar aspectos relacionados à qualidade, ao desempenho e à aplicação para a qual o computador vai ser orientado.

3.2 Componentes básicos de um computador

Apesar da existência de uma grande diversidade em termos de arquiteturas de computador, pode-se enumerar, num ponto de vista mais genérico os componentes básicos desta classe de equipamentos.

A Figura 1 apresenta um esquema de um computador, destacando os elementos que o compõem. Apesar da grande evolução ocorrida na área de informática desde o aparecimento dos primeiros computadores, o esquema apresentado na figura pode ser utilizado tanto para descrever um sistema computacional atual como os computadores da década de 40, projetados por engenheiros como John Von Neuman.

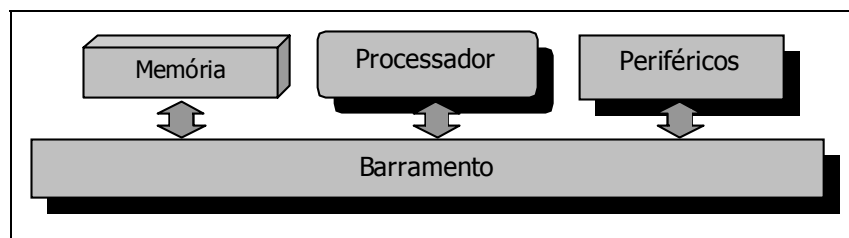


Figura 1. Elementos básicos do computador

Os principais elementos do computador são:

- O **processador** (ou microprocessador) é responsável pelo tratamento de informações armazenadas em memória (programas em código de máquina e dos dados).
- A **memória** é responsável pela armazenagem dos programas e dos dados.
- **Periféricos**, que são os dispositivos responsáveis pelas entradas e saídas de dados do computador, ou seja, pelas interações entre o computador e o mundo externo. Exemplos de periféricos são o monitor, teclados, *mouses*, impressoras, etc.
- **Barramento**, que liga todos estes componentes e é uma via de comunicação de alto desempenho por onde circulam os dados tratados pelo computador.

3.3 Processador

Um microprocessador, ou simplesmente processador, é um circuito integrado (ou chip), que é considerado o "cérebro" do computador (Figura 2). É ele que executa os

programas, faz os cálculos e toma as decisões, de acordo com as instruções armazenadas na memória.

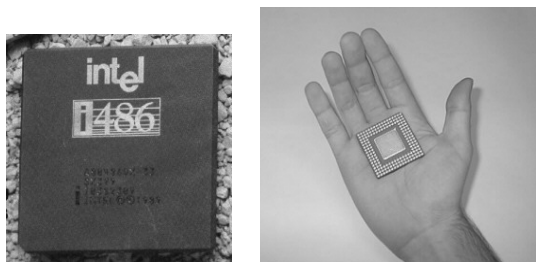


Figura 2. Microprocessador

Os microprocessadores formam uma parte importantíssima do computador, chamada de UCP (Unidade Central de Processamento), ou em inglês, CPU (*Central Processing Unit*). Antes da existência dos microprocessadores, as CPUs dos computadores eram formadas por um grande número de chips, distribuídos ao longo de uma ou diversas placas. Um microprocessador nada mais é que uma CPU inteira, dentro de um único CHIP. Podemos ver na Figura 3 um microprocessador e uma placa de circuito. Um microprocessador contém todos os circuitos que antigamente eram formados por diversas placas.

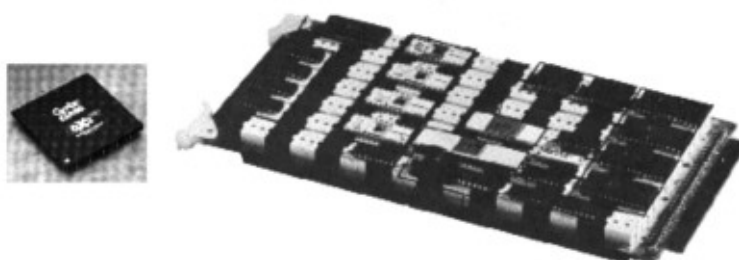


Figura 3. Microprocessador e placa de circuito impresso

Ligando-se um microprocessador a alguns chips de memória e alguns outros chips auxiliares, tornou-se possível construir um computador inteiro em uma única placa de circuito. Esta placa, como visto na Figura 4, é comumente chamada de placa mãe dos microcomputador.

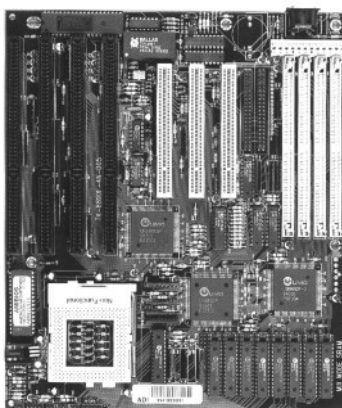


Figura 4. Placa Mãe

Não importa de que tipo de CPU estamos falando, seja um microprocessador, ou uma das várias placas que formam a CPU de um computador de grande porte, podemos dizer que a CPU realiza as seguintes tarefas:

- a) Busca e executa as instruções existentes na memória. Os programas e os dados que ficam gravados no disco (disco rígido ou disquetes), são transferidos para a memória. Uma vez estando na memória, a CPU pode executar os programas e processar os dados.

- b) Comanda todos os outros chips do computador.

A CPU é composta basicamente de três elementos: unidade de controle, unidade lógica e aritmética e registradores. As sessões que seguem apresentam cada um destes componentes.

3.3.1 Unidade Lógica e Aritmética

O primeiro componente essencial num computador (ou sistema computacional) é a **Unidade Lógica e Aritmética** (ALU), a qual, como o próprio nome indica, assume todas as tarefas relacionadas às operações lógicas (ou, e, negação, etc.) e aritméticas (adições, subtrações, etc...) a serem realizadas no contexto de uma tarefa.

Neste contexto, é importante observar a evolução que a ALU sofreu ao longo dos anos e quais são os parâmetros que influenciam no desempenho global de um sistema computacional:

- Um parâmetro importante é o **tamanho da palavra** processada pela unidade lógica e aritmética. Como o sistema de numeração adotado nas arquiteturas de computadores é o binário, o tamanho de palavra é dado em números de bits. Quanto maior o tamanho da palavra manipulada pelo microprocessador, maior é o seu potencial de cálculo e maior a precisão das operações realizadas.
- A **velocidade de cálculo** obviamente é outro fator de peso para o desempenho do computador, uma vez que ela será determinante para o tempo de resposta de um sistema computacional com respeito à execução de uma dada aplicação. A velocidade de cálculo está diretamente relacionada com a **frequência do relógio** que pilota o circuito da CPU como um todo.
- Outro parâmetro importante associado ao desempenho do computador é a quantidade de operações que ela suporta. Por exemplo, os primeiros processadores suportavam um conjunto relativamente modesto de operações lógicas e aritméticas. Em particular, no que diz respeito às operações aritméticas, os primeiros processadores suportavam apenas operações de adição e subtração, sendo que as demais operações tinham de ser implementadas através de seqüências destas operações básicas. Os processadores suportando um conjunto mais complexo de instruções surgiu de 15 anos para cá, graças à adoção da tecnologia CISC (*Complex Instruction Set Computer*).

A evolução nos tempos destes parâmetros será apresentada aqui a partir da história dos microprocessadores da família Intel e similares, resumida nas tabelas 1 e 2.

A sociedade Intel fundada em 1968, iniciou a fabricação de memórias para computadores, até que um fabricante de computadores lançou o desafio de construir uma unidade central de processamento (CPU), num único circuito para uma calculadora eletrônica. Assim, foi desenvolvida a CPU 4004 de 4 bits, e logo depois o 8008. Estes foram as primeiras CPUs integradas num único chip. A fabricação foi em pequena escala, pois a empresa não pensou que estas pudessem interessar a outros fabricantes. Esses chips eram caríssimos. Custavam, na época do seu lançamento, mais de 1000 dólares.

Em função do sucesso do 8008, a Intel lançou-se na fabricação de um novo chip que permitisse ultrapassar a barreira dos 16 Kbytes de memória, limite imposto devido ao número de pinos do 8008. Nasceu então o 8080 (1974), que provocou uma revolução no que diz respeito à indústria dos computadores. O 8080 foi o primeiro microprocessador a ser usado em larga escala nos chamados "computadores pessoais". Antes deles, os microcomputadores eram usados apenas em laboratórios científicos, em fábricas e em universidades. O 8080 popularizou o uso de microcomputadores por pequenas empresas e até para uso pessoal. Era comuns os micros pessoais baseados no 8080 e em outros microprocessadores rivais: o MC6800 da Motorola, o 6502, usando em um antigo microcomputador chamado de APPLE, e o Z-80 fabricado pela ZILOG, usado em um antigo computador chamado TRS-80. Surgia então a indústria dos microcomputadores. Ao mesmo tempo, surgia a indústria do software para microcomputadores, que criava programas de vários tipos para serem usados nessas máquinas. Os microcomputadores dessa época já tinham teclado, vídeo e impressora.

Seus dados e programas eram gravados normalmente em gravadores de fita K-7 adaptados para trabalhar com microcomputadores.

A INTEL produziu ainda, no final dos anos 70, um outro microprocessador para substituir o 8080. Chamava-se 8085. Todos esses microprocessadores (8080, 8085, Z-80, 6502, 6800 e outros) operavam com 8 bits. A Figura 5 mostra um dos primeiros microcomputadores brasileiros, o SCHUMEC M-101/85, lançado em 1981. Tinha um microprocessador INTEL 8085 de 6 MHz, 16 KB de memória e um gravador de fita K-7 para armazenamento de programas e dados. Seu monitor de vídeo era na verdade uma TV PHILIPS adaptada, já que nesta época o Brasil não fabricava monitores.

Tabela 1- Microprocessadores da Intel e suas características

Nome	Ano	Clock (MHz)	REG (bits).	Barramento (bits)	N. de Transistores	Memória Endereçável	Comentários
4004	1971	0,108	4	4	2.300	640 bytes	primeiro microprocessador (manipulação aritmética)
8008	1972	0,108	8	8	3.500	16 KB	primeiro micro 8 bits (manipulação de dados/caracteres)
8080	1974	2	8	8	6.000	64 KB	primeira CPU num chip (10x o desempenho do 8008)
8086	1978	5-10	16	16	29.000	1 MB	primeira CPU 16 bits num chip (10x o desempenho do 8086)
8088	1980	5-8	16	8	29.000	1 MB	processador do IBM-PC
80286	1982	8-12	16	16	134.000	16 MB	aumento no endereçamento (3-6x o desempenho do 8086)
80386DX	1985	16-33	32	32	275.000	4 GB	primeira CPU 32 bits
80386SX	1988	16-20	32	16	275.000	4 GB	80386 com barramento 80286
80486DX	1989	25-50	32	32	1.200.000	4 GB	versão 80386 mais rápida
80486 SX	1989	16-33	32	32	1.185.000	4 GB	80486 sem coprocessador matemático
Pentium	1993	60-166	32	32	3.100.000	4 GB	Arquitetura Super-escalar, 5x o desempenho do 486DX 33
Pentium Pro	1995	150-200	32	64	5.500.000	4 GB	Arquitetura de Execução dinâmica
Pentium II	1997	233-450	32	64	7.500.000	64 GB	Barramento Duplo independente, execução dinâmica, e tecnologia MMX
Pentium II Xeon	1998	400-450	32	64	7.500.000	64 GB	Para estações de trabalho e servidores
Pentium III	1999	450 a 1130	32	64	9.500.000	64 GB	PCs de negócio e de consumidores, servidores e estações de trabalho
Pentium III Xeon	1999	500 e 550	32	64	9.500.000	64 GB	PCs de negócio e servidores e estações de trabalho.

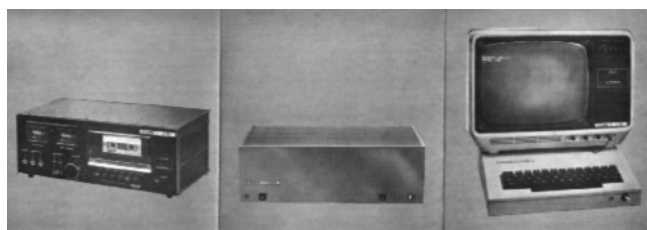


Figura 5. Microcomputador SCHUMEC M 101/85

Pouco depois, é lançada a primeira CPU em 16 bits, o 8086, concebido no mesmo espírito do 8080. O próximo passo foi o 8088, que possuía a mesma arquitetura do 8086, mas com um barramento interno de 8 bits e não de 16 como seu antecessor. Nessa ocasião, a IBM, que é o maior fabricante de computadores em todo o mundo, ainda não fabricava microcomputadores. Seus produtos eram os computadores de grande porte usados nos grandes centros de processamento de dados, e custavam alguns milhões de dólares. A IBM decidiu então entrar no mercado de computadores pessoais. Escolheu então o microprocessador 8088 para usar em seu microcomputador, chamado de "IBM Personal Computer", ou simplesmente, IBM PC. O IBM PC, o primeiro microcomputador de 16 bits, passou logo a dominar o mercado. Até os dias atuais, os modernos microcomputadores são compatíveis com o IBM PC original, lançado em 1981.

Tabela 2- Microprocessadores da Intel e AMD e suas características

Type/Generation	Year	Data/ Address bus width	Level 1 Cache (KB)	Memory <u>bus speed</u> (MHz)	Internal clock speed (MHz)
8088/First	1979	8/20 bit	None	4.77-8	4.77-8
8086/First	1978	16/20 bit	None	4.77-8	4.77-8
80286/Second	1982	16/24 bit	None	6-20	6-20
80386DX/Third	1985	32/32 bit	None	16-33	16-33
80386SX/Third	1988	16/32 bit	8	16-33	16-33
80486DX/Fourth	1989	32/32 bit	8	25-50	25-50
80486SX/Fourth	1989	32/32 bit	8	25-50	25-50
80486DX2/Fourth	1992	32/32 bit	8	25-40	50-80
80486DX4/Fourth	1994	32/32 bit	8+8	25-40	75-120
<u>Pentium</u> /Fifth	1993	64/32 bit	8+8	60-66	60-200
MMX/Fifth	1997	64/32 bit	16+16	66	166-233
<u>Pentium Pro</u> /Sixth	1995	64/32 bit	8+8	66	150-200
<u>Pentium II</u> /Sixth	1997	64/32 bit	16+16	66	233-300
Pentium II/Sixth	1998	64/32 bit	16+16	66/100	300-450
Pentium III/Sixth	1999	64/32 bit	16+16	100	450-1.2GHz
AMD Athlon/Seventh	1999	64/32 bit	64+64	266	500-2.2GHz
Pentium 4/Seventh	2000	64/32 bit	12+8	400	1.4GHz-3.6GHz
AMD Athlon 64/Eighth	2003	64/64 bit	64+64	400	2GHz-2.4GHz

A grande vantagem do IBM PC-XT em relação ao IBM PC era a possibilidade de operar com um disco rígido (também chamado de winchester) de 10 MB, uma altíssima capacidade para aquela época.

Os microprocessadores 80186 e 80188 foram extensões aos dois chips anteriores, do ponto de vista das E/S. Entretanto, estes nunca foram chips altamente utilizados.

O passo seguinte e natural seria a passagem para 32 bits, o que foi concretizado com o lançamento do 80386 e 80386SX. Logo após surgiu o 80486 e que possui, integrados, um coprocessador de ponto flutuante e um controlador de memória.

Atualmente, os computadores são baseados nos processadores Pentium II e III, cujas características estão na tabela acima.

3.3.2 Unidade de Controle (UC)

A Unidade de Controle tem a maior importância na operação de um computador, uma vez que é esta unidade que assume toda a tarefa de controle das ações a serem realizadas pelo computador, comandando todos os demais componentes de sua arquitetura. É este elemento quem deve garantir a correta execução dos programas e a utilização dos dados corretos nas operações que os manipulam. É a unidade de controle que gerencia todos os eventos associados à operação do computador, particularmente as chamadas **interrupções**, tão utilizadas nos sistemas há muito tempo.

3.3.3 Registradores

A CPU contém internamente uma memória de alta velocidade que permite o armazenamento de valores intermediários ou informações de comando. Esta memória é composta de registradores (ou registros), na qual cada registro tem uma função própria. Os registros, geralmente numerosos, são utilizados para assegurar o armazenamento temporário de informações importantes para o processamento de uma dada instrução. Conceitualmente, registro e memória são semelhantes: a localização, a capacidade de armazenamento e os tempos de acesso às informações que os diferenciam. Os registros se localizam no interior de um microprocessador, enquanto a memória é externa a este. Um registro memoriza um número limitado de bits, geralmente uma palavra de memória. Os registros mais importantes são:

- Contador de programa (PC - *Program Counter*), que aponta para a próxima instrução a executar.
- Registro de instrução (IR - *Instruction Register*) que armazena a instrução em execução.
- Outros registros que permitem o armazenamento de resultados intermediários.

3.3.4 Clock

Clock é um circuito oscilador que tem a função de sincronizar e ditar a medida de velocidade de transferência de dados no computador, por exemplo, entre o processador e a memória principal. Esta frequência é medida em ciclos por segundo, ou Hertz.

Existe a frequência própria do processador, comandando operações internas a ele, e a frequência do computador a ele associado, basicamente ciclos CPU-Memória principal.

Os processadores Pentium-100, Pentium MMX-233, Pentium II-300, acessam a memória principal a 66 MHz. Suas frequências respectivas de 100, 233 e 300 MHz são atingidas, tão somente, no interior do chip. Dizem, portanto, respeito ao processamento interno do processador e não à frequência na relação CPU-Memória do computador.

Já os processadores Pentium II-350 e superiores tem uma frequência externa de 100 MHz, acarretando um desempenho melhor do microcomputador, tanto no processamento propriamente dito quanto nas operações de disco e vídeo.

3.4 Processadores CISC e RISC

Uma instrução em um programa de alto nível é implementado por diversas instruções de processador a mais baixo nível. Por exemplo, uma instrução de um programa que imprime um conjunto de caracteres na tela é realizado a nível de processador por um conjunto de instruções.

De acordo com o número de instruções suportadas, pode-se classificar um processador como RISC (Reduced Instruction Set Computing) ou CISC (Complex Instruction Set Computing).

Os microprocessadores são, em sua maioria, chips CISC. Um processador CISC suporta um conjunto maior de instruções, sendo cada instrução mais especializada; graças a isso, ele pode executar, de modo direto, a maioria das operações programadas pelos programas de alto nível. Devido a isto, o número de instruções de processador para implementar uma instrução de mais alto nível são menores. Mas quanto maior a quantidade de instruções que um processador suporta, mais lenta é a execução de cada uma delas.

Um processador RISC reconhece um número limitado de instruções, que, em contrapartida, são otimizadas para que sejam executadas com mais rapidez. Portanto, a arquitetura RISC reduz o conjunto de instruções ao mínimo, e as instruções não implementadas diretamente são realizadas por uma combinação de instruções existentes. Mas em compensação, um programa é implementado por um número maior de instruções.

A Apple desenvolveu, sob a arquitetura RISC, em parceria com a Motorola e a IBM, o chip Powerpc, com mais poder de processamento que o chip Pentium. Ele equipa a linha de microprocessadores MacPower (nos modelos 6100, 7100, 8100).

3.5 Memória

Todo computador é dotado de uma quantidade de memória (que pode variar de máquina para máquina) a qual se constitui de um conjunto de circuitos capazes de armazenar os dados e os programas a serem executados pela máquina. Nós podemos identificar diferentes categorias de memória:

- a **memória principal**, ou memória de trabalho, onde normalmente devem estar armazenados os programas e dados a serem manipulados pelo processador;
- a **memória secundária** que permitem armazenar uma maior quantidade de dados e instruções por um período de tempo mais longo; o disco rígido é o exemplo mais evidente de memória secundária de um computador, mas podem ser citados outros dispositivos menos recentes como as unidades de fita magnética e os cartões perfurados;
- a **memória cache**, que se constitui de uma pequena porção de memória com curto tempo de resposta, normalmente integrada aos processadores e que permite incrementar o desempenho durante a execução de um programa.

Os circuitos de memória são normalmente subdivididos em pequenas unidades de armazenamento, geralmente um byte. Cada uma desta unidade é identificada no circuito por um endereço único, o qual vai ser referenciado pelo processador no momento de consultar ou alterar o seu conteúdo. Por exemplo, no caso do processador 8088 nós temos um espaço de endereçamento de 1 MBytes, então este endereço único vai de 0 a FFFFF, conforme ilustração abaixo. As quantidades de memória hoje são definidas em termos de Kbytes (quilobytes) que correspondem a 1024 bytes ou (2^{10} bytes) e MBytes (megabytes), que correspondem a 1024 KBytes ou (2^{20} bytes).

Endereço hexadecimal	Conteúdo de memória
00000	00
00001	23
00002	00
...	...
...	...
FFFFE	E4
FFFFF	FF

3.5.1 Memória Principal

A **memória principal**, ou memória de trabalho, é onde normalmente devem estar armazenados os programas e dados a serem manipulados pelo processador. Geralmente é esta memória que se referencia na especificação de um microcomputador. E hoje, as quantidades de memória mais usuais disponíveis nos microcomputadores são 32, 64, 128 e 256 MBytes de memória, ou mesmo até mais para aplicações específicas.

Em termos básicos, a memória principal é vista como um conjunto de chips que são inseridas na placa mãe do computador. Pode-se ver na Figura 6 alguns tipos de chips de memória, usados para formar a memória principal de microcomputadores. Os dois módulos superiores desta figura são as memórias utilizadas atualmente. Mais adiante serão apresentadas as tecnologias de memória e os tipos de empacotamento.

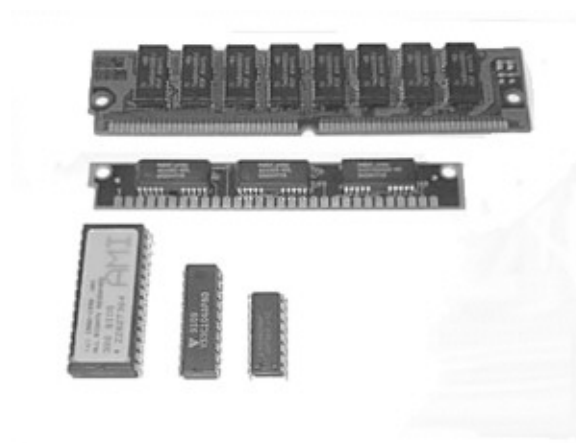


Figura 6. Chips de memória

Os módulos de memória mais usados nos microcomputadores são os dois módulos de memória superiores da Figura 6. O mais superior é conectado no barramento PCI e o segundo no barramento ISA (estes tipos de barramento serão apresentados mais adiante). Estes módulos são colocados na placa mãe, como apresentado na Figura 7.

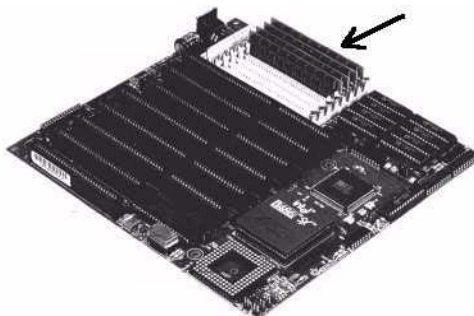


Figura 7. Placa Mãe e suas memórias

Tipos de memória

Os chips de memória podem ser divididos em duas grandes categorias:

- **RAM** (memória de leitura e escrita): são chips de memória que podem ser lidos e gravados pela CPU a qualquer instante. A CPU usa a RAM para armazenar e

executar programas vindos do disco, para ler e gravar os dados que estão sendo processados. Uma outra característica da RAM, é que se trata de uma memória VOLÁTIL. Isso significa que quando o computador é desligado, todos os seus dados são apagados. Por essa razão, é necessário que os programas e dados fiquem gravados no disco, que é uma memória PERMANENTE.

Existem vários tipos de RAM com diversas características e para diversas aplicações. A mais conhecida é a DRAM (dinâmica) e a SRAM (estática) e suas evoluções. Estes tipos serão detalhados mais adiante.

- **Memórias Não Voláteis:** são memórias cujas informações mantidas não são perdidas caso o computador seja desligado. Nos microcomputadores, existe um programa muito importante chamado de BIOS (*Basic Input-Output System* - Sistema Básico de Entrada e Saída). O BIOS tem várias funções, entre as quais, a de realizar a "partida" do computador. Quando ligamos o computador, o BIOS realiza a contagem de memória, faz uma rápida checagem do funcionamento do computador e realiza a carga do Sistema Operacional que deve estar armazenado no disco. O BIOS está gravado em uma memória permanente localizada na placa mãe. Existem diversos tipos de memória permanente:
 - **ROM:** São chips de memória que podem ser lidos pela CPU a qualquer instante, mas não podem ser gravados pela CPU. Sua gravação é feita apenas pelo fabricante do computador, ou pelo fabricante de memórias. Os dados armazenados nela já saem prontos de fábrica e são produzidos em larga escala na indústria. A característica importante de ROM é que trata-se de uma memória PERMANENTE. Seu conteúdo nunca é perdido, mesmo com o computador desligado. Portanto este tipo de memória é usada para armazenar programas estáticos (que não alteram) e produzidos em massa. Este tipo de memória foi usado para armazenar o BIOS, que se localiza na placa-mãe.
 - **PROM:** Significa Programmable ROM, ou seja, ROM programável. Trata-se de uma espécie de ROM que é produzida apagada. O fabricante pode programá-las, ou seja, gravar seu programa. Esta gravação pode ser feita apenas um vez, pois utiliza um processo irreversível. Por isso, usa-se o termo queimar a PROM quando se grava nesta memória.
 - **EPROM:** Significa Erasable PROM, ou seja, uma ROM programável e apagável. Assim como ocorre com a PROM, a EPROM pode ser programada e a partir daí, comporta-se como uma ROM comum, mantendo os dados armazenados mesmo sem corrente elétrica, e permitindo apenas operações de leitura. A grande diferença é que a EPROM pode ser apagada com raios ultravioleta de alta potência. Possuem uma "janela de vidro", através da qual os raios ultravioleta podem incidir nas operações de apagamento. Nota-se que essa janela de vidro fica sempre coberta por um adesivo que tampa a passagem de luz. É fácil identificar um chip EPROM na placa mãe justamente pela presença desse adesivo.
 - **EEPROM:** Significa Electrically Erasable Programmable ROM (EEPROM ou E2PROM). Esta é o tipo de memória ROM mais flexível, que pode ser apagada sob o controle de software. Este é o tipo que se usa para armazenar as BIOS atuais. Dessa forma, o usuário pode realizar atualizações no BIOS, fornecidas pelo fabricante da placa de CPU. Quando se ouve falar em "flash BIOS" ou "fazendo um upgrade de BIOS", isto se refere a reprogramação do BIOS EEPROM com um programa de software especial.

Memória fora da placa mãe

Como já mostramos, a placa mãe contém quase toda a memória de um microcomputador, mas outras placas também podem conter memórias, do tipo RAM e do tipo ROM. Por exemplo, as placas de vídeo contém uma ROM com o seu próprio BIOS, e contém uma RAM chamada de memória de vídeo, que armazena os caracteres e gráficos que são mostrados na tela. Podemos ver na Figura 8 os chips de memória existentes na PLACA DE VÍDEO de um microcomputador.

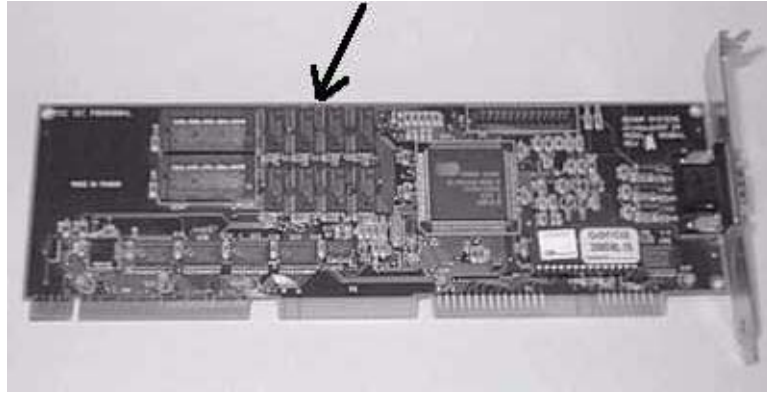


Figura 8. Placa de vídeo e suas memórias

3.5.2 Tecnologias de memória RAM

Como já apresentado, existem vários tipos de RAM com diversas características e para diversas aplicações. A mais conhecida é a SRAM (estática) e as várias tecnologias DRAM (dinâmica).

SRAM

A Static Ram é um tipo de RAM que mantém seus dados sem uma reatualização externa (refresh), na medida que seus circuitos são alimentados continuamente. Ao contrário, as DRAM necessita de pulsos de 15ns para manter seu conteúdo, de forma que a energia não fique o tempo todo abastecendo os chip. Esse pulso periódico é o refresh.

Toda a memória RAM é mais lenta ou mais rápida de acordo com o tempo de acesso medido em nano-segundos (bilionésimos de segundo). A SRAM é mais rápida que a DRAM. É comum hoje encontrar nas DRAMs 6 a 150ns de tempo de acesso (dependendo da tecnologia DRAM), enquanto que nas SRAM é de até 15ns. Mas ela apresenta algumas fraquezas. SRAM tem um custo por byte muitas vezes maior que as DRAM. Cada bit da DRAM necessita de um transistor e de um capacitor (que quando energizado mantém a carga elétrica se o bit contém um “1” ou sem carga se ele contém um 0”). Já a memória SRAM não necessita de refresh, como se a corrente elétrica estivesse o tempo todo ligada a ela, por meio de interruptores que acionam e fecham a memória. Assim, o tempo de acesso é menor. Mas a sua desvantagem é o seu tamanho, pois requer seis transistores para cada bit.

Estas vantagens e desvantagens mostram claramente que em termos de desempenho as SRAM são melhores que DRAM. A única coisa que proíbe o uso generalizado da memória SRAM é seu custo. Por isto, as SRAMs são usadas basicamente para compor a memória cache (visto mais adiante).

Memórias DRAM

Como apresentado, a DRAM é um tipo de memória RAM que apenas mantém os dados se eles são continuamente reforçados por um circuito lógico especial chamado circuito refresh. Devido a esta construção, também a ação de leitura refresca os conteúdos da memória. Se isto não for feito regularmente, então a DRAM perderá seus conteúdos, mesmo se a alimentação for mantida. Esta ação de refrescamento é que dá o nome a memória de dinâmica.

Todos os PCs usam DRAM para constituir sua memória de sistema (em vez de SRAM). A razão é simples, elas são mais baratas e tomam menos espaço, tipicamente $\frac{1}{4}$ da área de silício das SRAMs ou menos. Existem várias espécies de tecnologias de DRAM e as velocidades que elas fornecem são diferentes. Estas tecnologias serão vistas mais adiante.

Os chips de DRAM diferenciam nos seguintes aspectos:

- número de células na memória;
- tamanho de cada célula na memória;

- tempo de acesso;
- encapsulamento;

O número de células é relacionado com a capacidade de armazenamento (ou posições de memória), existindo chips com 8KB a 16MB de células de memória. O tamanho das células de memória é o número de bits que cada célula armazena. Existem chips de memória com 1, 4, 8, 9, 32 ou 36 bits. O tempo de acesso, como já dito, é em bilionésimos de segundo, conhecidos como nano-segundos.

Até o final dos anos 80, a memória DRAM era feita com o encapsulamento DIP, que tinha que ser encaixada na placa-mãe. Logo depois surgiu o encapsulamento SIPP, que deu lugar, por sua vez, ao encapsulamento SIMM. Veja cada um dos chips abaixo.

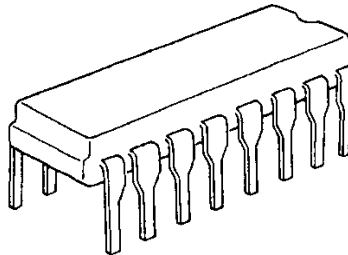


Figura 9. Um chip de memória DIP

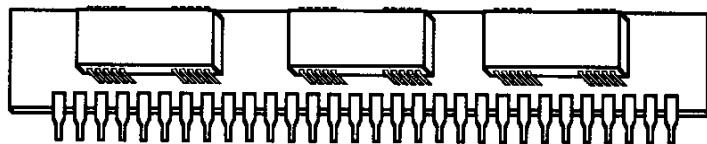


Figura 10. Um módulo de memória SIPP



Figura 11. Um módulo de memória SIMM

Até hoje os chips de memória que compõem as placas adaptadoras são do tipo DIP (*Dual In-Line Package*).

Com o SIPP (*Single In-Line Pin package*), surgiu o que é chamado módulos de memória, que eram vários chips de DRAM numa fileira de terminais onde se encaixavam num soquete. Esse tipo de encapsulamento foi bastante usado até o início dos anos 90.

O SIMM (*Single In-Line Memory Module*) surgiu ao redor de 1992. Ela é eletricamente igual aos SIPP, possuindo de diferente apenas a forma de seus contatos para afiação na placa-mãe. Podemos dizer que o SIPP possui perninhas e o SIMM, contatos na borda inferior.

Entre 1992 e 1994, usou-se muito os módulos de memória SIMM pequenos, de 30 pinos. Esses módulos operavam com 8 bits cada um, necessitando serem usados em grupos. Então, para completar um banco de memória num 80386, era necessário 4 desses módulos, pois 4x8bits significa 32 bits, o suficiente para um 80386 ou 80486. Nota-se que os módulos de 30 pinos (ou vias) conseguiam compor no máximo 4MB em um único módulo.

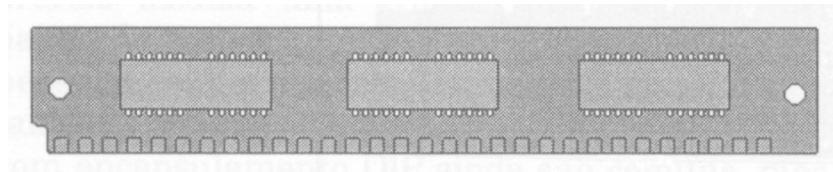


Figura 12. Módulo de memória de 30 vias

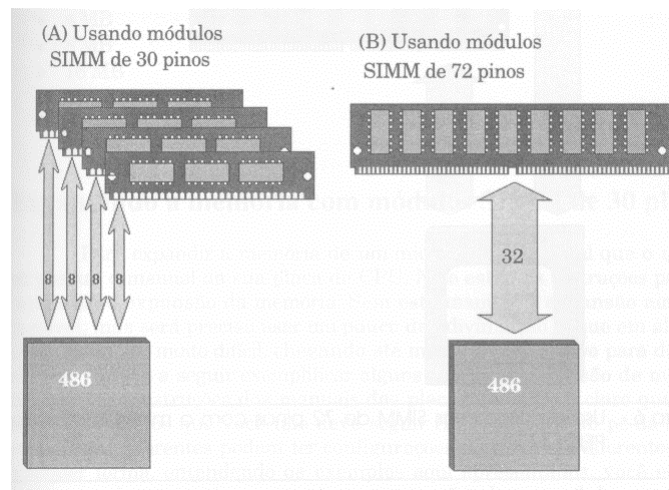


Figura 13. Combinação para chegar a 32 bits num 486

Mais tarde, surgiu os módulos SIMM de memória de 72 vias, que é um pouco maior que o de 30, operando a 32 bits, onde os últimos 486 fabricados usavam muito (algumas vezes em conjunto com os de 30 vias). Esses módulos de memória de 72 vias podem ter até 32MB em um único módulo. Como esses módulos são de 32 bits, para poder completar um banco num Pentium, que é de 64 bits, são necessários 2 módulos.

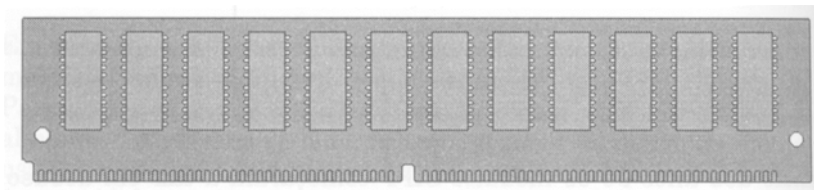


Figura 14. Módulo de memória de 72 vias:

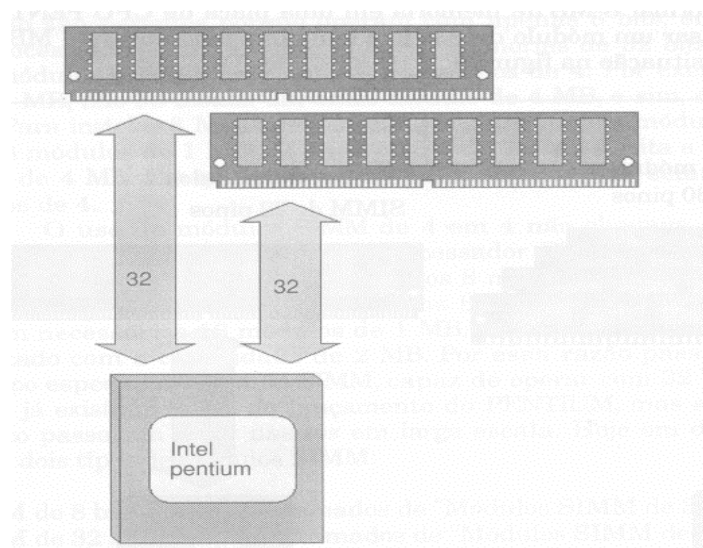


Figura 15. Combinação para se chegar a 64 bits num Pentium

Logo, a partir de 1995, praticamente todas as placas-mãe passaram a ser fabricadas com soquetes que suportavam esses módulos de 72 vias. Em outros casos, as placas-mãe podiam possuir ainda um banco de 30 vias e dois bancos de 72 vias, isto é, quatro soquetes de 30 vias e 2 soquetes de 72 vias.

E em 1997 surgiram as memória no encapsulamento DIMM (Dual In-Line Memory Module), que é um módulo de memória com um encaixe igual ao do SIMM, mas que é de 168 pinos, praticamente o dobro do tamanho de um SIMM. Essa memória é de 64 bits. Assim, para um Pentium, basta um desses módulos de memória para funcionar.

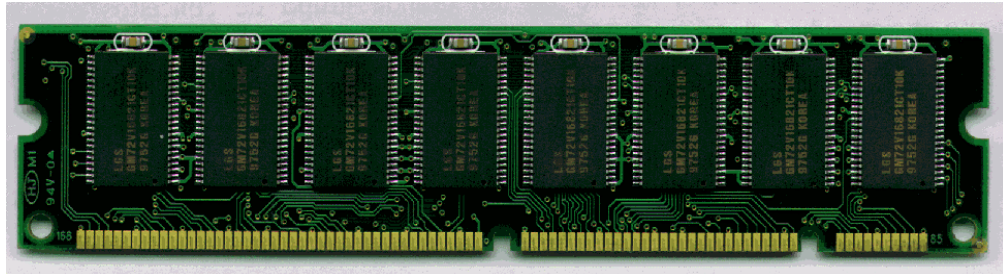


Figura 16. Módulo de memória DIMM de 32MB

É normal encontrar nos computadores de hoje 32 a 256MB de memória, podendo haver mais que isto. Para cada quantidade, deve-se ter a combinação certa de memória a ser colocada. O número de barramentos de memória da placa mãe define a quantidade máxima de memória de um computador.

Asynchronous and Synchronous DRAM

A DRAM convencional, do tipo que era usada nos PCs desde os dias dos IBM PCs originais, é dita ser assíncrona. Isto se refere ao fato que a memória não é sincronizada com o relógio do sistema (clock). Um acesso a memória é iniciado, e um certo tempo posterior o valor de memória aparece no barramento. Os sinais não são coordenados com o clock do sistema. Estas memórias assíncronas trabalham bem com barramentos de memória de baixa velocidade, mas não trabalham bem com sistemas de memória de alta velocidade (>66MHz).

O novo tipo de DRAM, chamada "synchronous DRAM" ou "SDRAM", é sincronizada com o relógio do sistema, todos os sinais são ligados ao clock de maneira que ela é melhor controlada. Como clock é coordenado pela clock da CPU, o tempo de execução de comandos e transmissão de dados é reduzido. Este tipo de memória é muito mais rápida que a DRAM assíncrona e pode ser usada para melhorar o desempenho do sistema. Portanto as SDRAM são melhores adaptadas aos sistemas de memória de alta velocidade dos novos PCs.

Praticamente todos os micros novos são vendidos com um tipo de memória chamada SDRAM, mas é sempre bom confirmar com o vendedor se a sua memória é desse tipo ou de uma tecnologia mais antiga (como EDO ou FPM), que devem ser evitadas atualmente, pois encontram-se obsoletas e são mais lentas que as memórias SDRAM.

Memórias PC-66 e PC-100

Existem dois tipos de memória SDRAM no mercado atualmente: PC-66 e PC-100. O primeiro tipo deve ser utilizado por processadores que trabalham externamente a 66 MHz, como os processadores Pentium II até 333 MHz e Celeron. Já o segundo tipo deve ser obrigatoriamente utilizado pelos processadores que operam a 100 MHz externamente, como o Pentium II a partir de 350 MHz, Pentium III, K6-2 a partir de 300 MHz e K6-III.

As memórias SDRAM de 10ns de tempo de acesso são, teoricamente, capazes de trabalhar a 100 MHz. Mas, na prática, foi constatado que esse tipo de memória era instável para sistemas de 100 MHz. Com isso, somente são aconselhadas para micros com barramento de 100 MHz memórias com tempo de acesso de 8 ns ou menos. Essas memórias são as PC-100.

Caso uma memória PC-66 for utilizada em processadores que operam externamente a 100 MHz (em vez de memórias PC-100), o computador apresenta diversos problemas de congelamentos e resets aleatórios, além dos conhecidos erros de tela azul da morte e GPF ("Esse programa executou uma operação ilegal e será fechado"). É por esse motivo que, ao fazer um upgrade de um processador com barramento de 66 MHz para um com barramento de 100 MHz, é necessário trocar também a memória RAM.

Identificar memórias PC-66 e PC-100 é relativamente simples, são aquelas que tem um tempo de acesso de 8ns ou inferior. Toda memória possui estampada, em seus circuitos, uma codificação após um sinal de "-". Memórias -10 ou com valores superiores (-12, -15, etc) são memórias do tipo PC-66. Já memórias com marcação -8 ou inferiores (-7, etc.) são memórias do tipo PC-100.

DDR or SDRAM II

Double-data rate SDRAM é uma versão mais rápida de SDRAM que é capaz de ler dados na subida e descida do clock do sistema, assim dobrando a taxa de dados do chip de memória.

Rambus DRAM® (Rambus™ RDRAM)

RDRAM, desenvolvida pela empresa Rambus, é uma memória extremamente rápida e usada como um canal rápido para transmitir dados a velocidades de 10 vezes mais rápidas que as DRAM padrões. Espera-se que esta tecnologia seja usada como memória principal dos PCs. A Figura 17 mostra alguns módulos mais modernos.

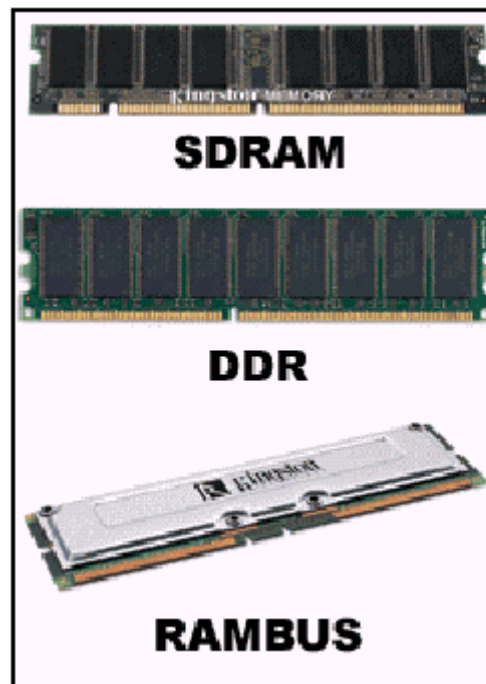


Figura 17 – Módulos de memória RAM¹

3.5.3 Memória Secundária

Além da memória principal, que é diretamente acessada pela CPU, existe também a memória secundária (também chamada de memória de massa). Este tipo de memória não é acessada diretamente pela CPU. Seu acesso é feito através de interfaces ou controladoras especiais. A memória secundária é uma memória do tipo permanente (não

¹ <http://www.computermemoryupgrade.net/computer-memory-types-and-memory-technology.html>

se apaga quando o computador está desligado), que tem uma alta capacidade de armazenamento, e um custo muito mais baixo que o da memória principal.

A memória secundária não é formada por chips, e sim, por dispositivos que utilizam outras tecnologias de armazenamento. A Figura 17 apresenta alguns exemplos de memória secundária: o disco rígido, disquetes, CD-ROM e fita magnética. O disco rígido, assim como os disquetes e as unidades de fita, usam a tecnologia magnética para armazenar dados. Os discos CD-ROM usam tecnologia ótica.



Figura 17. Dispositivos de armazenamento secundário

Arquivos

Arquivo nada mais é que um conjunto de dados gravados na memória secundária (disco rígido, disquete, fita magnética, CD-ROM, etc). Os arquivos são uma forma de organizar melhor os dados dentro da memória secundária. Nos microcomputadores, os arquivos normalmente recebem nomes, como por exemplo: CURRÍCULO.DOC, COMMAND.COM, CHKDSK.EXE. O nome completo de um arquivo é dividido em duas partes, separadas por um ponto. A primeira parte é chamada de NOME do arquivo. A segunda parte é chamada de extensão

Os arquivos podem armazenar diversos tipos de dados:

- **Instruções para a CPU:** Dizemos que se trata-se um **arquivo executável**. Os arquivos executáveis, nos microcomputadores, normalmente usam a extensão COM ou EXE. Por exemplo: FORMAT.COM, COMMAND.COM, SORT.EXE, XCOPY.EXE
- **Documentos:** São textos digitados com o auxílio de um tipo de programa chamado de editor de textos, ou processador de textos. Normalmente esses arquivos usam a extensão TXT ou DOC. Por exemplo: CURRIC.DOC, LISTA.TXT, REUNIAO.DOC
- **Gráficos e Imagens:** São arquivos que representam figuras. Essas figuras podem ser vistas na tela ou na impressora, com o auxílio de programas apropriados. Por exemplo, foto.jpg, logo.gif, fundo.bmp.
- **Dados genéricos:** que são os demais tipos de arquivos, por exemplo, arquivos de som e vídeo.

Muitas vezes os programas precisam manipular uma quantidade de dados tão grande que não cabem na memória principal. Nesse caso, esses dados são armazenados em arquivos que são lidos da memória secundária e processados por partes. Muitas vezes esses dados podem até caber na memória principal, mas por uma questão de organização ficam armazenados em arquivos.

Dados armazenados em memória secundária podem ser lidos e transformados em impulsos eletrônicos e enviados para a CPU e para a memória. A CPU, por sua vez, pode acessar os dados originários dos arquivos e enviá-los para os dispositivos de saída, como o monitor e a impressora, para que possamos vê-los com nossos próprios olhos. Por exemplo, se um arquivo contém uma carta, só poderemos ver esta carta depois que o arquivo for lido pela CPU e enviado para a impressora.

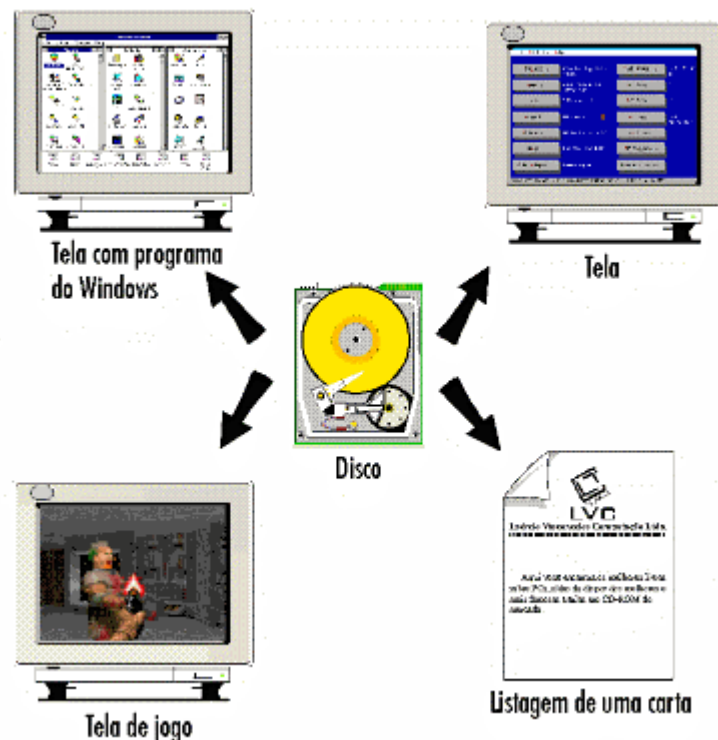


Figura 18. Arquivos "vistos" na tela e na impressora

Discos Rígidos

Em termos gerais, um disco rígido usa discos achatados chamados pratos, revestido nos dois lados por um material magnético projetado para armazenar informações. Os pratos são montado em uma pilha. Estes pratos (o disco) giram a uma rotação constante (3600 a 7200 rpm) desde que o computador é ligado. Dispositivos especiais de leitura/escrita, chamados de cabeçotes, são usados para escrever ou ler informações no/do disco, sendo que sua posição no disco é controlada por um braço atuador. Cada prato contém dois cabeçotes, um na parte superior do prato e outro na parte inferior. Assim, um disco rígido com dois pratos tem quatro cabeçotes. Todos os cabeçotes são presos a um único braço atuador, eles não se movem individualmente.

Os discos rígidos e a CPU se comunicam via um conjunto de circuitos denominados controladora de disco rígido que está geralmente integrada na placa-mãe, mas pode ser uma placa independente. Um exemplo de placa controladora é a SCSI (Small Compact System Interface), que permite a conexão de diversos periféricos, inclusive de naturezas distintas. Outra controladora muito usada é a IDE (Intelligent Drive Electronics).

Dados são organizados no disco em cilindros, trilhas e setores (Figura 19). É a formatação que marca o disco com trilhas e setores. Cilindros são trilhas concêntricas na superfície dos discos. Hoje, existem aproximadamente 3000 trilhas em cada lado de um prato de 3,5 polegadas. Uma trilha é dividida em setores. Cada setor tem o tamanho de 512 bytes. Em uma operação de leitura de um setor, o controlador de disco interpreta o endereço do dado e move os cabeçotes para o cilindro que contém os dados. Quando os cabeçotes estão na posição correta, o controlador ativa o cabeçote específico para ler a trilha que contém o dado. O cabeçote então lê a trilha procurando o setor que contém o dado para leitura. A placa controladora do disco coordena o fluxo de informação vinda do disco rígido em uma área de armazenamento temporária (*cache*). Ela então envia a informação pela interface do disco rígido.

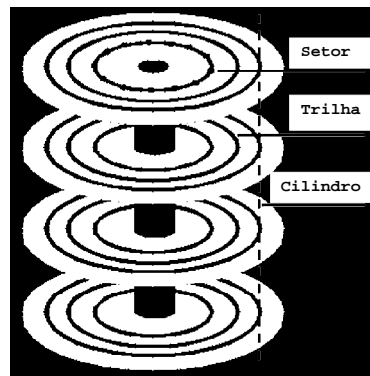


Figura 19. Disco Rígido

Muitas vezes os blocos de dados de um arquivo estão espalhados ao redor do disco. Portanto, alguns mecanismos são necessários para rastrear os blocos de um arquivo (p.e. lista ligada, FAT – *File Allocation Table* no DOS, I-node do UNIX). Quando da leitura de vários blocos em um arquivo espalhado, uma busca deve ser realizada para a leitura de cada bloco.

Os XTs usavam discos rígidos com 10 MB, mas atualmente existem discos rígidos com várias capacidades. Em 1991, a maioria dos microcomputadores usava discos rígidos com 40, 60 ou 80 MB. Já em 1993, os discos de 120 e 200 MB eram mais comuns. Atualmente a capacidade dos discos rígidos mais usuais são: 4.3, 6.4, 8.4, 9.1, 10.2 GB. Mas existem discos rígidos com maior capacidade, como 18.2 GB.

Considere por exemplo, um computador com 64 MB de RAM (que faz parte da memória principal) e um disco rígido de 8.4 GB (que faz parte da memória secundária). Os 8.4 GB do disco rígido servem para armazenar diversos programas e diversos dados. Os 64 MB de RAM servem para manter apenas os programas e os dados que estão sendo efetivamente processados. Podemos também mostrar como o custo da memória secundária é muito mais baixo que o da memória principal. Enquanto 64 MB de RAM custa cerca de R\$ 250,00 (preço de janeiro/2000), um disco rígido de 8.4 GB custa cerca de R\$ 420 (preço de janeiro/2000).

CD-ROM (Compact-Disk Read-Only Memory)

O CD-ROM é o meio de armazenamento que tem o menor custo por cada MB armazenado. Infelizmente, como o nome já diz, esse tipo de disco não pode ser usado para gravar dados. Pode ser usado apenas para leitura. O CD-ROM tem sido utilizado para veiculação de software, dicionários, enciclopédias, etc.

Um CD-ROM utiliza técnicas óticas de laser em vez de eletromagnetismo. A leitura de um CD é feita com a emissão de um feixe de laser sobre a superfície do disco.

Sob a mesma tecnologia, encontram-se os WORM (Write Once, Read Many), discos óticos que podem ser gravados apenas uma vez, mas lidos inúmeras vezes, e os discos apagáveis (Magnet Optical Erasable Disk), regraváveis, que permitem inúmeras atualizações.

Unidades de fita magnéticas

A unidade de fita magnética é o dispositivo de alta velocidade que lê e grava fitas magnéticas. A fita magnética, geralmente acondicionada em cartuchos, é o meio de armazenamento de grande capacidade. Um rolo pode conter centenas de megabytes de dados. Elas são normalmente utilizadas para backups de discos rígidos, também pode ser utilizadas no processamento em que os dados tenham que ser acessados seqüencialmente, principalmente quando todo o arquivo ou grande parte dele tem que ser acessado. Seu uso é mais corrente junto a máquinas de porte médio ou grande.

Memória secundária como Memória principal

A memória secundária é muito mais barata, de maior capacidade, e ainda é permanente, ou seja, não apaga os dados quando o computador é desligado. Por que então esse tipo

de memória não é usado no lugar da memória principal? Já pensou como seria interessante acoplar um disco rígido de 8.4 GB diretamente a um microprocessador? Você iria dizer: "... possuo um computador Pentium III 450 com 8.4 GB de memória...". Infelizmente, os dispositivos de armazenamento secundário são extremamente lentos. São lentos demais para serem acoplados diretamente ao microprocessador. Além disso, não permitem acessos a seus bytes individuais, como um microprocessador precisa realizar. Os meios de armazenamento secundário só permitem o acesso a bloco de dados. Em um disquete ou em um disco rígido, por exemplo, as leituras são feitas em unidades mínimas chamadas de setores. Cada setor tem 512 bytes. Para ter acesso a um único byte, é preciso ler o setor inteiro. Isso faz com que seu acoplamento direto à CPU seja inviável. As memórias RAM e ROM são milhares de vezes mais rápidas e permitem que sejam feitos acessos a qualquer um de seus bytes, de forma individual. Por isso são usadas para formar a memória principal.

Os meios de armazenamento secundário são mais lentos porque envolvem movimento de suas partes mecânicas. O disco rígido, o drive de disquete e o drive de CD-ROM precisam mover a cabeça de leitura até o ponto onde será feita a leitura. A unidade de fita precisa girar a fita até o ponto a ser acessado. As memórias ROM e RAM não precisam desses movimentos, pois seu acesso é inteiramente eletrônico. Por isso são usadas como memória principal.

3.5.4 Memória Cache

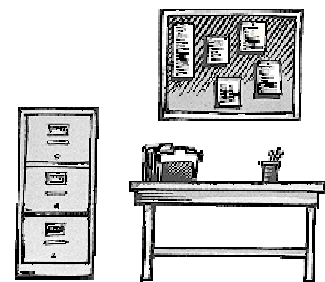
Memória cache é uma área reservada de memória que possui duas funções: aumentar o desempenho do computador e aumentar o tempo de vida das unidades de disco. Basicamente há dois tipos de memória cache:

- A que vem incorporada à máquina.
- A que é implementada via software na memória RAM do sistema.

A memória cache incorporada à máquina é um tipo muito mais rápido do que a memória RAM convencional. Por exemplo, em uma placa mãe típica de 100MHz, o tempo de acesso à memória principal é de 180 nanosegundos, e o acesso à memória cache é na ordem de 45 nanosegundos. Por isso ela é usada para armazenar tabelas muito usadas pelo sistema operacional ou para executar parte de programas que necessitam de maior velocidade de processamento.

A memória cache criada via software é usada para aumentar o desempenho do acesso ao disco do sistema, guardando as informações mais acessadas na memória, ou seja, quando for preciso acessar uma nova informação, ela já está armazenada em memória, que possui um tempo de acesso muito mais rápido do que o disco. Com isso, o número de vezes que a unidade de disco é acessada diminui, reduzindo o desgaste físico do disco e da cabeça de leitura e gravação.

Poderia-se fazer uma analogia entre a memória cache e o fichário que ficaria em nossa mesa de trabalho. Um arquivo (memória principal) maior conteria informações completas para realização do trabalho e o fichário (a memória cache) conteria informações mais corriqueiras. Mais próximo do trabalhador (a CPU), o fichário (a cache) aumentaria a rapidez na realização do trabalho e reduziria a pesquisa no arquivo.



Níveis de Cache

Existem vários "níveis" de cache em um computador moderno, cada nível agindo como um buffer para informações recentemente usadas para aumentar o desempenho. Mas quando referimos apenas a simplesmente "cache", normalmente está se referenciando o nível "secundário" ou nível 2 que é aquela posicionada entre o processador e a memória principal.

Cada nível é mais próximo do processador e mais rápida que o nível mais abaixo. Cada nível também cacheia o nível mais abaixo dela, devido a sua velocidade aumentada relativa aos níveis mais baixo.

Cache Primária – Layer 1

A cache nível 1 (L1) ou cache primária é a memória mais rápida de um PC. Ela é de fato construída diretamente no processador. Esta cache é muito pequena, geralmente de 9KB a 64KB, mas é extremamente rápida. Ela roda na mesma velocidade do processador. Se o processador solicita uma informação e pode encontrá-la na cache L1, este é o melhor caso, pois a informação é imediatamente disponível e o processador não tem que esperar.

O processador Pentium III apresenta uma memória cache L1 de 16KB para instruções e outra de 16KB para dados.

Cache Secundária – Layer 2

A cache nível 2 é uma cache secundária para o nível 2, e é maior e um pouco mais lenta. Ela é usada para armazenar endereços recentes que não são mantidos pelo nível 1. Este nível tem um tamanho variando de 64KB a 2MB e até pouco tempo atrás era encontrada na placa mãe ou um módulo que é inserido na placa mãe.

Os processadores Pentium Pro têm a cache L2 no mesmo pacote do processador (embora este não esteja no mesmo circuito do processador e da cache L1). Isto significa que ela roda muito mais rápida que a cache L2 que é separada e reside na placa mãe. O

Pentium II e III até 600MHz têm um meio termo, sua cache L2, de 512KB, roda na metade da velocidade da CPU. O Pentium III acima 600MHz apresenta uma memória cache L2 de 256KB no mesmo pacote e opera na mesma velocidade da CPU.

O nível dois algumas vezes é chamada de cache externa, desde que ela reside fora do processador. Mesmo nos Pentium Pro, que está em um chip separado no mesmo pacote.

Cache Externas

As primeiras placas-mãe equipadas com SRAM apresentavam pequenas quantidades de cache, entre 8KB a 128KB. Com o 486, tornou-se mais comum que a placa-mãe tivesse 128KB a 512KB de cache. O Pentium geralmente tem 512KB a 1MB de cache.

Até recentemente, os chips que formavam a SRAM eram modularizados em forma de chips DIP, mas atualmente, fabrica-se os módulos do tipo COAST, semelhante ao SIMM, porém, eletronicamente diferentes.



Figura 20. Módulo COAST.

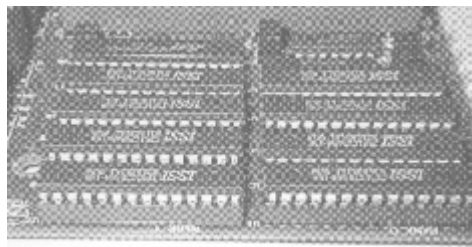


Figura 21. SRAM do tipo DIP:

Praticamente, não se vê mais placas de CPU novas com memória cache do tipo DIP. As placas de CPU atuais incorporam os chips de SRAM na placa mãe, mas não do tipo DIP. Os chips são soldados e geralmente não há bancos vazios para a expansão. Com os módulos COAST, basta trocar o módulo. E com os chips VLSI da Intel, estes possuem a capacidade de detectar automaticamente o tipo de cache em uso, através do BIOS, dispensando a configuração manual, de forma que logo no processo de boot há essa identificação.

Nos processadores Intel mais modernos, a partir do Pentium II a *cache* L2 fica embutida no *chip* (*on-chip*).

3.6 Barramentos

Um barramento, ou *bus*, nada mais é do que um caminho comum pelo qual os dados trafegam dentro do computador. Este caminho é usado para comunicações e pode ser estabelecido entre dois ou mais elementos do computador.

O tamanho de um barramento é importante pois ele determina quantos dados podem ser transmitidos em uma única vez. Por exemplo, um barramento de 16 bits pode transmitir 16 bits de dado, e um barramento de 32 bits pode transmitir 32 bits de dados a cada vez.

Todo barramento tem uma velocidade medida em MHz. Um barramento rápido permite transmitir dados rapidamente, que tornam as aplicações mais rápidas.

Um PC tem muitos tipos de barramentos, que incluem:

- **Barramento do processador:** é o barramento que o chipset usa para enviar/receber informações do processador. O chipset são os chips de suporte adjacentes contidos na placa mãe.
- **Barramento de Cache:** usado pelos Pentium Pro e Pentium III, é um barramento dedicado para acessar o sistema cache. Ele é algumas vezes chamado de barramento *backside*. Processadores convencionais que usam placas mãe de quinta geração tem uma cache conectada ao barramento de memória padrão.
- **Barramento de memória:** é um barramento que conecta o sub-sistema de memória ao chipset e ao processador. Em alguns sistemas o barramento do processador e o barramento de memória são basicamente a mesma coisa.
- **Barramento local de E/S** (Entrada/Saída). É usado para conectar periféricos de alto desempenho à memória, chipset e processador. Por exemplo, placas de vídeo, interface de redes de alta velocidade geralmente usam um barramento deste tipo. Os dois tipos barramentos locais de E/S mais comuns são o VESA Local Bus (VLB) e o Peripheral Component Interconnect Bus (PCI).
- **Barramento padrão de E/S:** conecta os três barramentos acima ao antigo barramento padrão de E/S, usado para periféricos lentos (modems, placas de som regulares, interfaces de rede de baixa velocidade) e também para compatibilidade com dispositivos antigos. O principal barramento padrão de E/S é o Industry Standard Architecture (ISA) bus. Nos PCs, o barramento ISA (antigo padrão) está sendo substituído por barramentos mais rápidos, como PCI. Geralmente, quando se ouve sobre o barramento de E/S, praticamente falamos de slots de expansão.

Alguns novos PCs usam um “barramento” adicional projetado para comunicação de grafismos apenas. A palavra “barramento” foi colocada entre aspas pois esta tecnologia não se trata de um barramento, ela é um porto: o Accelerated Graphics Port (AGP). A diferença entre um porto e um barramento é que o barramento é geralmente projetado para vários dispositivos compartilhando um meio de comunicação, enquanto um porto é apenas entre dois dispositivos.

Todos os barramentos tem duas partes: um barramento de endereçamento e um barramento de dados. O barramento de dados transfere o dado em si (por exemplo, o valor de memória) e o barramento de endereço transfere a informação de onde o dado se encontra.

3.6.1 Barramento do processador

O barramento do processador é o caminho de comunicação entre a CPU e o chipset. Este barramento é usado para transferir dados entre a CPU e o barramento principal do sistema (Figura 22).

Pelo motivo de que o barramento do processador serve para conseguir informações para e da CPU à velocidade mais rápida possível, este barramento opera de forma muito mais rápida que qualquer outro tipo de barramento no computador. Não existem

congestionamentos neste barramento. Este barramento consiste em circuitos elétricos de dados, endereçamento e controle. No processador 80486, o barramento consiste em 32 linhas de endereçamento, 32 linhas de dados e algumas linhas para controle. No Pentium, são 32 linhas de endereçamento, 64 linhas de dados e as linhas de controle.

O barramento do processador opera a mesma velocidade que a CPU opera externamente. Num Pentium 100MHz, o barramento opera a 66,6MHz, bem como num Pentium 133MHz, 166MHz e 200MHz. No processador Pentium II, o barramento de processador opera a 100MHz e no Pentium III, a 100 ou 133 MHz.

O barramento do processador é ligado nos pinos de conexão externos do processador e pode transferir bits de dados a cada um ou dois ciclos. Um 80486 transfere 32 bits de dados por vez, enquanto um Pentium transfere 64 bits.

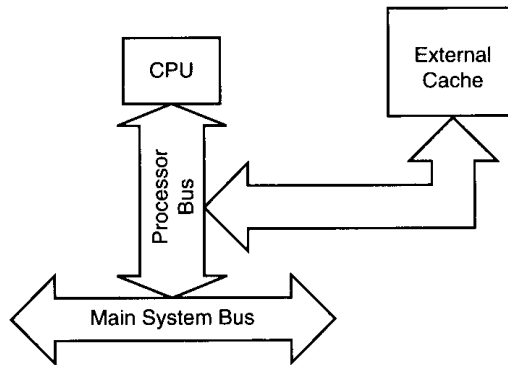


Figura 22. Funcionamento o barramento do processador

3.6.2 Barramento de Cache

Processadores convencionais usam a cache externa na placa mãe e a CPU acessa esta cache usando o barramento padrão de memória. Para obter um melhor desempenho, a maioria dos novos usam um barramento de alta velocidade conectando o processador à cache L2. No caso do Pentium Pro e Pentium III acima de 600MHz, este barramento é interno ao processador e opera na taxa da CPU. No caso dos Pentium II e III até 600MHz, o barramento de cache opera na metade da taxa do processador. Por exemplo, um Pentium II 266 MHz tem um barramento de cache operando a 133 MHz.

3.6.3 Barramento da Memória

O barramento da memória é usado para transferir informação entre a CPU e a memória principal do sistema (a memória RAM). Estes barramento é uma parte do barramento do processador, ou na maioria dos casos é implementado separadamente por um chipset dedicado que é responsável pela transferência de informação entre o barramento do processador e o barramento da memória (Figura 23).

Devido aos limites da capacidade do chips de DRAM, a taxa de transferência de informação no barramento da memória é menor que a do barramento do processador.

Vale lembrar que a largura do barramento da memória é a mesma que do barramento do processador. Igualmente, num processador Pentium ou Pentium Pro, que possui um barramento de 64 bit, cada banco de memória que é usado deverá ter 64 bits preenchidos. As placas-mãe atuais possuem um suporte a pentes de memória do tipo DIMM, que são de 64 bits. Assim, somente um pente é necessário para o funcionamento do computador.

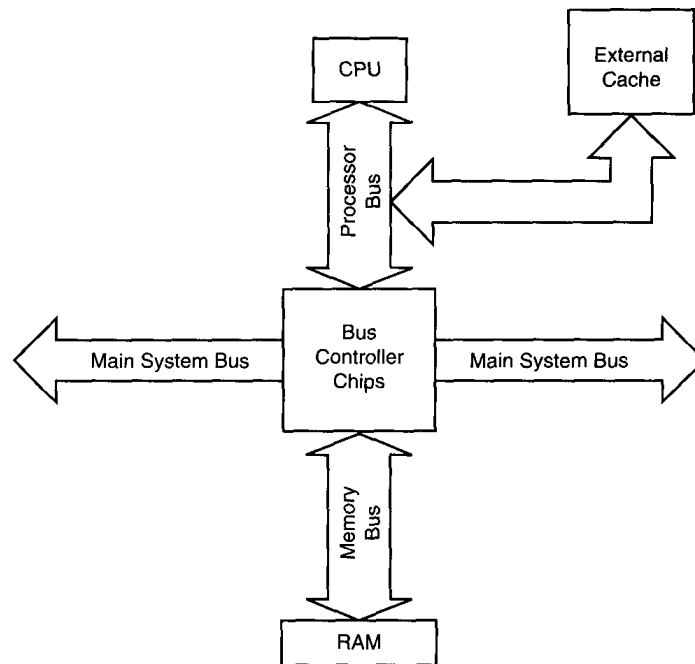


Figura 23. Funcionamento do barramento da memória

3.6.4 Barramento de E/S

Desde a criação do PC, muitos barramento de E/S tem sido desenvolvidos, pela simples razão de CPUs mais rápidas, maiores demandas de software e maiores requisitos de vídeo. Para cada avanço obtido nos computadores, requer-se um barramento que atenda às exigências do hardware. Mesmo assim, necessita-se que seja padronizado diante do mercado a nível mundial, sem grandes custos aos usuários (esta também é uma das razões em que os barramento de E/S não puderam evoluir mais do que se esperava). Até foram criados barramento de E/S mais rápidos do que o mercado exige, mas não foram colocados com sucesso no mercado justamente pelos custos gerados na atualização de hardware por parte dos usuários).

Pode-se encontrar estes barramento de E/S na arquitetura do computadores de hoje: ISA, MCA, EISA, VESA Local Bus, PCI Local Bus.

3.6.5 Barramento ISA

Este barramento foi introduzido no PC original em 1981 como um barramento de 8 bits e mais tarde foi expandido para 16 bits com o PC/AT em 1984, chamando-se arquitetura padrão da indústria. O padrão ISA é ainda é utilizado em arquiteturas menos recentes, mas este tipo de barramento é geralmente substituído pelo PCI. Mas o padrão ISA ainda é rápido o bastante para muitos periféricos que usamos hoje, como uma placa fax/modem ou uma placa de som.

Existem duas versões do barramento ISA:

- O barramento ISA de 8 bits surgiu com o XT e opera a 4,77MHz. Esta arquitetura de barramento foi usada inicialmente nos primeiros PCs e é desnecessária nos sistemas atuais. O slot ISA de 8 bits é um conector com 62 contatos, possuindo eletronicamente 8 linhas de dados e 20 linhas de endereçamento, habilitando ao slot manusear 1MB de memória (Figura 24).



Figura 24. Slot ISA de 8 bits

- O barramento ISA de 16 bits consegue operar num máximo de 8,33MHz, com uma taxa máxima de 8MB/s. Surgiu após a produção de processadores de 16 bits (a partir do 80286, em 1984). Para compatibilidade, a IBM teve que criar um tipo de slot que suportasse ambos 8 bits e 16 bits. Criou-se um slot com duas divisões (Figura 25). As placas de expansão de 8 bits poderiam ser colocadas na parte maior do slot e as de 16 bits em ambas as partes do slot. O conector ISA de 16 bits passa a ter, assim, mais 36 pinos de conexão em relação ao ISA de 8 bits, em função do amplo caminho de dados. Esse padrão de slots é usado ainda hoje nos computadores Pentium. Todas as placas de CPU possuem três ou quatro slots em cada placa.

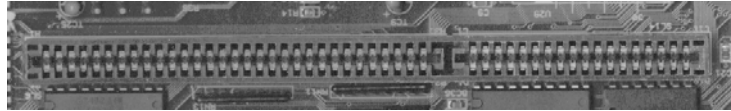


Figura 25. Slot ISA de 16 bits

3.6.6 Barramento MCA

Com a introdução de chips de 32 bits (80386DX), significou que o ISA não poderia suportar todo o poder dessa nova geração de CPUs. Com o lançamento do 80386DX, a IBM decidiu criar um novo barramento, que resultou no MCA, arquitetura de micro canal. Porém, a IBM não quis somente criar um novo barramento, mas também receber os royalties por ele de outros fabricantes. Outra razão para o MCA não se tornar um padrão da indústria foi que não suportava as placas ISA de 16 bits e 8 bits, haja vista que o conector era fisicamente menor, mais complexo e bem mais diferente.

3.6.7 Barramento EISA

Este barramento foi criado em 1988 em resposta ao barramento MCA, sendo primeiramente fabricado pela Compaq. Mesmo assim, poucas placas adaptadoras foram criadas a partir do EISA. Simplesmente não emplacou no mercado, assim como a MCA. Porém, o que salvou esse padrão foi o de poder usar as placas ISA presentes. Isto é, os slots EISA estão presentes, mas foram usados principalmente pelas placas ISA, que não necessitaram dos recursos colocados no EISA.

3.6.8 Local bus

Os barramentos de E/S vistos até agora (ISA, MCA e EISA) possuem algo em comum: velocidade baixa. Esta limitação vem dos PCs originais, onde o barramento de E/S operava a mesma velocidade do barramento do processador. Como a velocidade do processador aumentou, o barramento de E/S aumentou apenas em largura da banda de dados, permanecendo a uma velocidade muito mais baixa, por causa da enorme base de placas adaptadoras instaladas que só poderiam operar a velocidades mais baixas (Figura 26).

A idéia de que o sistema rodasse mais lento do que deveria, iria insatisfazer os usuários. Porém, a velocidade do barramento de E/S muitas vezes é só um incômodo, porque não será necessário altas velocidades para a comunicação com o teclado ou mouse, por exemplo. O problema só existe na lentidão do barramento quando placas de vídeo e controladores de disco se tornam prejudicados. Com os sistemas e ambientes operacionais gráficos, tanto o vídeo como os discos necessitam de mais velocidade para os dados. Muito mais do que os 8MHz oferecidos com o ISA.

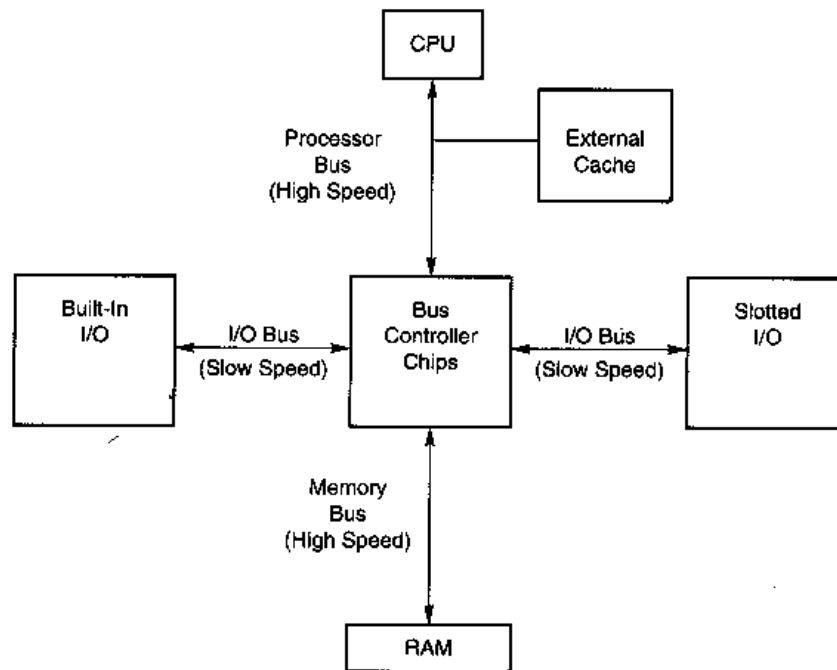


Figura 26. Layout tradicional do barramento dos PCs

Uma solução óbvia é mover os caminhos de dados dos slots de E/S para a área onde pode acessar velocidades maiores como a do barramento do processador, como um cache externo (Figura 27). Este arranjo torna-se conhecido como *local bus*, já que as placas adaptadoras podem acessar diretamente o barramento do processador, que é local. Conseqüentemente, para comportar esse acesso, os slots precisam ser diferentes do que os existentes, bem como para diferenciar as placas mais lentas das mais rápidas que são destinadas a esse tipo de barramento.

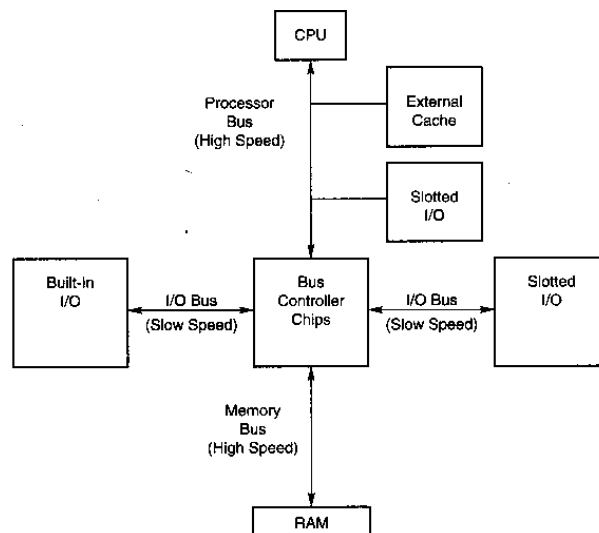


Figura 27. Funcionamento do barramento local

Nota-se que os primeiros barramento ISA de 8 e 16 bits até rodavam a velocidade do barramento do processador, usando este barramento local. Mas quando o processador se tornou mais rápido (e seu barramento também), o barramento principal do ISA teve que ser desacoplado do barramento do processador.

3.6.9 VESA Local Bus

VESA Local Bus (VLB) foi o primeiro barramento local popular, em 1992. VESA é abreviação de *Video Electronics Standards Association*, que foi formada por um grupo de fabricantes interessados em problemas relacionados com vídeos em computadores pessoais. A principal razão do desenvolvimento do VLB foi aumentar o desempenho de vídeo nos PCs.

O VLB é um barramento de 32 bits que é de certo modo uma extensão do barramento processador/memória do 486. Um slot VLB é um slot ISA 16 bits com um terceiro e quarto slots no final. Desta maneira, uma placa ISA pode ser conectada no VLB.

O VLB roda a taxas de 33 MHz, embora velocidades maiores são possíveis para alguns sistemas.

O VLB foi muito popular durante a era 486, com a introdução do Pentium e seu barramento local PCI, em 1994, o VLB começou a ser abandonado. Além das vantagens do PCI em relação ao VLB, existem também várias problemas do VLB que contribuíram para seu abandono:

- Seu projeto foi duramente baseado no processador 486, e a adaptação para o Pentium causou sérios problemas de compatibilidade e outros problemas.
- O barramento em si é eletricamente fraco, por exemplo, o número de placas que poderiam ser usadas no barramento era pequeno (normalmente duas ou mesmo uma), e ocorria problemas quando mais de uma placa era conectada.
- O barramento não tinha um controle muito forte, desde que não havia um bom esquema de árbitro, e não suporta o Plug and Play.

3.6.10 Barramento PCI

PCI é a abreviação de *Peripheral Component Interconnect*, um padrão de barramento local desenvolvido pela Intel Corporation. Embora ele tenha sido desenvolvido pela Intel, o PCI não é ligado a qualquer família de microprocessadores.

A maior parte dos PCs modernos incluem um barramento PCI em adição ao barramento de expansão ISA. Mas a tendência é que o barramento ISA seja totalmente substituído pelo PCI.

Como o VESA Local Bus, PCI é um barramento de 32 bits que normalmente roda a um máximo de 33 MHz. Nestas condições, ele funciona a 133 MB/s. A vantagem chave do PCI sobre seu predecessor, o VLB, é a existência de um circuito no chipset que controla o barramento. O VLB era basicamente uma extensão do barramento do processador 486. O PCI deste modo não é casado com 486, e seu chipset fornece funcionalidades para controle, que habilita o PCI a fazer mais coisas que o VLB poderia fazer.

A especificação 2.1 do barramento PCI estende o barramento para 64 bits e opera com a velocidade de 66MHz. Assim, a velocidade pode ser quadruplicada. Na prática o barramento PCI de 64 bits ainda não foi implementado no PC (existem em outras plataformas como o Digital Equipment's Alpha). Atualmente a velocidade é limitada a 33 MHz na maior parte dos PCs.

A operação do barramento PCI pode ser configurado como síncrona ou assíncrona, dependendo do chipset e da placa mãe. Na configuração sincronizada (usado por maior parte dos PCs), o PCI roda na metade da velocidade do barramento de memória; como este barramento opera usualmente a 50, 60 ou 66 MHz, o PCI poderia rodar a 25, 30 ou 33 MHz, respectivamente. No modo assíncrono, a velocidade do PCI pode ser setado independente da velocidade do barramento de memória. Isto é controlado por jumpers na placa mãe, ou setado pela BIOS.

O barramento PCI opera concorrentemente com o barramento do processador. A CPU pode processar os dados como um cache externo enquanto o barramento PCI está ocupado transferindo informação entre outras partes do sistema.

Além do ser mais eficiente, o barramento PCI incorpora o recurso Plug and Play, não necessitando que o usuário configure as placas adaptadoras. O sistema se encarrega

disso. Para isso, os computadores devem possuir um BIOS com essa característica. Por exemplo, o Award Plug and Play BIOS possui esse recurso.

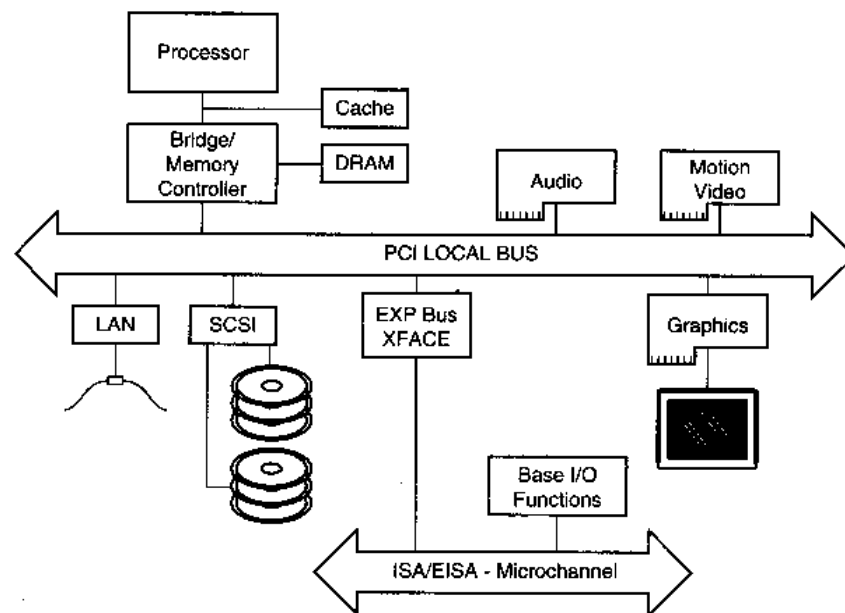


Figura 28. Estrutura do barramento PCI

Os dados trafegam através do barramento PCI a 33MHz, com toda a banda de dados da CPU, isto é, usando-se uma CPU de 32 bits, a largura é de 32 bits, com uma taxa de transferência de 132MB por segundo no máximo.



Figura 29. Slot PCI de 32 bits

O slot PCI de 64 bits é constituído de uma divisão a mais, como nos VLB. Com esta banda de dados, a velocidade dobrando para um máximo de 264MB/s.

3.6.11 Barramento AGP

A grande demanda por velocidade obrigou a indústria da informática a desenvolver um barramento específico para aplicações gráficas, o *Accelerated Graphics Port* (AGP). O barramento AGP permite que a placa de vídeo acesse a memória RAM diretamente. Este recurso é muito utilizado nas placas 3D.

3.6.12 Barramento PC-Card (PCMCIA)

Este barramento é destinado aos computadores portáteis (notebooks e laptops), desenvolvido pela PCMCIA, com um conjunto de mais de 300 fabricantes, onde se estabeleceu os padrões para os cartões adaptadores e de expansão para notebooks e laptops.

Um cartão PCMCIA usa de 68 conectores, sendo ainda mais caro que o ISA. Apesar dos padrões, a indústria flexibilizou demais a arquitetura, de forma que alguns cartões podem não ser compatíveis com algum equipamento ou outro. Os cartões também possuem o recurso de ser Plug and Play.

3.6.13 USB

A interface externa USB fornece uma comunicação serial de 12 Mbps, apenas sobre uma conexão de 4 fios. Uma única porta USB pode ser usado para conectar até 127 periféricos, tal como *mouse*, modems, teclados, *scanners*, câmeras.

A USB também atende às especificações *Plug and Play* da Intel, inclusive de poder conectar os dispositivos com a máquina ligada e sem precisar reiniciá-las. Simplesmente se conecta o dispositivo e a USB irá detectar automaticamente e alocar os recursos necessários para o seu funcionamento. O suporte a USB é fornecido com o Windows 95 e mais ainda pelo BIOS.

Outra característica importante é que os dispositivos USB conectados têm o fornecimento de energia pelo próprio USB, o que significa que o USB verifica quais os requisitos de energia requeridos pelos periféricos e avisar se estes dispositivos excederem os limites.

A USB 1.1 transfere dados na velocidade entre 1,5 e 12 Mbps (Megabits por segundo) , já a USB 2.0 chega a taxa de 480 Mbps.

3.6.14 Firewire (IEEE 1394)

O firewire é um barramento serial padrão externamente rápido que suporta taxas de transferências de dados de até 400 MBps. Um único porto 1394 pode ser usado para conectar até 63 dispositivos externos. Além da alta velocidade, o 1394 também suporta dados isócronos, transmitindo dados com uma taxa garantida. Isto é ideal para dispositivos que necessitam altas taxas de transferências em tempo-real, tal como dispositivos de vídeo.

Embora muito rápido, o 1394 é muito caro. Tal como o USB, o 1394 suporta Plug-and-Play, e também fornece potências aos periféricos. A principal diferença entre do 1394 e USB é que o 1394 suporta taxas de transferências altas e é mais caro. Por esta razão, é esperado usado mais freqüentemente para dispositivos que necessitam taxas altas, como câmeras de vídeo, e o USB conectará os demais periféricos.

3.7 Dispositivos de Entrada/Saída

Os Dispositivos de Entrada/Saída são equipamentos utilizados como portadores das informação que o computador irá processar. Por exemplo, quando se pressiona uma tecla, faz com que o teclado transmita o código da tecla pressionada. Este código é recebido por um circuito chamado de INTERFACE DE TECLADO. Ao receber o código de uma tecla, a interface de teclado avisa a CPU que existe um caractere recebido. Por outro lado, quando a CPU precisa enviar uma mensagem para o usuário, precisa que a mensagem seja colocada na tela. Isto é feito com auxílio de um circuito chamado de INTERFACE DE VÍDEO. A CPU envia para a interface de vídeo, a mensagem, seja ela em forma de texto ou figura. A interface de vídeo coloca então a mensagem na tela.

Existem portanto no computador, os chamados DISPOSITIVOS DE ENTRADA E SAÍDA, também chamados de PERIFÉRICOS. Através desses dispositivos, o computador pode armazenar, ler, transmitir e receber dados. Dentre os diversos dispositivos de E/S, existem alguns que são especializados apenas em ENTRADA, outros especializados apenas em SAÍDA e outros em ENTRADA E SAÍDA. Podemos citar os seguintes exemplos:

- ENTRADA:
 - Teclado - Lê os caracteres digitados pelo usuário
 - MOUSE - Lê os movimentos e toque de botões
 - Drive de CD-ROM - Lê dados de discos CD-ROM
 - Microfone - Transmite sons para o computador
 - SCANNER - Usado para o computador "ler" figuras ou fotos
- SAÍDA:

- Vídeo - Mostra ao usuário, na tela caracteres e gráficos
- Impressora - Imprime caracteres e gráficos
- Alto-falante - Realiza comunicação com o usuário através de som
- ENTRADA E SAÍDA:
 - Disco rígido - Grava e lê dados
 - Drive de disquete - Grava e lê dados em disquetes
 - Unidade de fita magnética - Grava e lê dados em fitas magnéticas
 - MODEM - Transmite e recebe dados pela linha telefônica

As Figura 30, 32 e 33 mostram diversos tipos de periféricos.



Figura 30. Dispositivos de ENTRADA



Figura 31. Dispositivos de SAÍDA

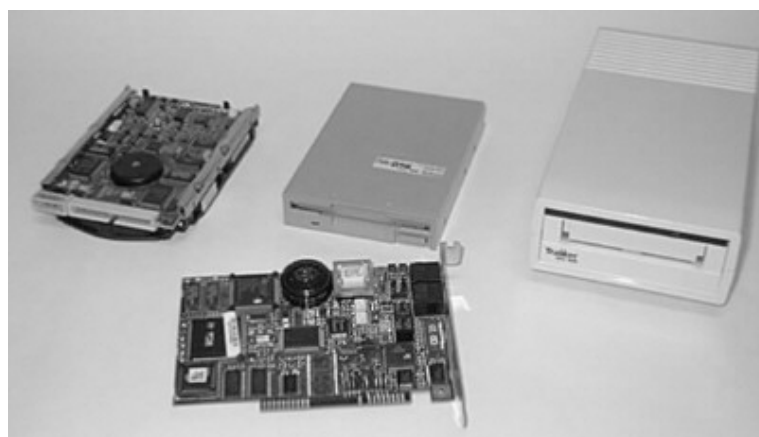


Figura 32. Dispositivos de ENTRADA E SAÍDA

Os microcomputadores muitas vezes são vendidos com apenas alguns de seus periféricos de uso mais prioritário. Nesse caso, dizemos que trata-se de uma CONFIGURAÇÃO BÁSICA. Um exemplo tipo de configuração básica é aquela em que o microcomputador é acompanhado de:

- Monitor

- Drives
- Disco rígido
- Teclado
- Mouse
- Alto falante interno

O usuário pode, nesse caso, adquirir novos periféricos e realizar sua instalação no microcomputador. Essa tarefa pode ser realizada por usuários mais experientes, ou então por técnicos especializados. Quando instalamos novos periféricos em um computador, dizemos que estamos realizando um EXPANSÃO. Nesse caso, podem ser instalados, por exemplo:

- Impressora
- Kit multimídia (alto falantes, placa de som e drive de CD-ROM)
- Modem
- SCANNER

Tipos de comunicação com os Dispositivos

A CPU não pode comunicar-se diretamente com os periféricos. Esta comunicação é feita com a ajuda de circuitos chamados de interfaces ou portas de E/S, as quais podem implementar a transmissão das palavras de dados segundo duas diferentes políticas:

- a **comunicação paralela**, onde cada dígito (ou bit) da palavra de dados é conduzido por um fio dedicado, o que significa que os cabos utilizados para a comunicação paralela são dotados de uma grande quantidade de fios (ou vias); o exemplo mais clássico de dispositivos que utilizam a comunicação paralela são as impressoras;
- a **comunicação serial**, onde os bits de cada palavra são transmitidos um a um, de forma seqüencial, através de uma única via, o que explica o fato dos cabos que implementam este tipo de comunicação serem constituídos por uma pequena quantidade de fios; os exemplos mais conhecidos de dispositivos que fazem uso desta política de comunicação são o mouse e os modems.