

[< VOLTAR](#)

Visão geral do sistema de armazenamento e hierarquia de memória

Conhecer os dispositivos de armazenamento por meio do conceito e dos tipos de memórias utilizadas no computador.

[NESTE TÓPICO](#)[Marcar tópico](#)

Subsistemas de memória

Memória é o dispositivo físico capaz de armazenar algum tipo de conteúdo para uso no computador. Esse conteúdo pode ser:

- Instruções
- Dados
- Resultados de processamento intermediário
- Resultados de processamento final (informação)

Apesar de parecer simples como conceito, a memória de um computador exhibe, talvez, a mais vasta gama de: tipos, tecnologia, organização, rendimento e custos, dentre todas as especificidades de um sistema de computação. Nenhuma tecnologia pode ser considerada ótima em termos da satisfação dos requisitos de memória de um sistema de computação. Como consequência, o sistema de computação típico está equipado com uma hierarquia de subsistemas de memória, alguns internos ao sistema (diretamente acessíveis ao processador) e outros externos (acessíveis ao processador por meio de um módulo de E/S).

Memória interna e memória externa

A memória interna (também chamada de primária) é responsável pelo armazenamento temporário de dados utilizados no processamento. É essencial para que ocorra o processamento. Corresponde aos registradores,



cache e memória principal.

Já a memória externa (secundária) armazena dados não voláteis em dispositivos periféricos acessíveis por meio de controladores de E/S. Alguns exemplos: disco rígido, CD-rom etc.

Na base da pirâmide que representa a hierarquia (apresentada em item posterior) de memória em um sistema de computação encontra-se um tipo de memória com maior capacidade de armazenamento do que os outros tipos já descritos, menor custo por byte armazenado e com tempos de acesso também superiores aos outros tipos. Essa memória, denominada memória secundária, memória auxiliar ou memória de massa, tem por objetivo garantir um armazenamento mais permanente a toda a estrutura de dados e programas do usuário, razão pela qual deve naturalmente possuir maior capacidade que a memória principal.

A memória secundária de um sistema de computação pode ser constituída por diferentes tipos de dispositivos, alguns diretamente ligados ao sistema para acesso imediato (discos rígidos, por exemplo), e outros que podem ser conectados quando desejado (como os disquetes, fitas de armazenamento, CD-ROM etc.), cuja informação armazenada se torna diretamente conectada e disponível para o específico disquete ou fita que estiver inserido no elemento de leitura/escrita (drive ou acionador), enquanto os demais ficam disponíveis (off-line) para acesso manual pelo usuário.

Capacidades

Por causa da simplicidade de projeto e construção, acarretando a redução de seu custo e maior confiabilidade, os circuitos eletrônicos que formam os computadores digitais atuais são capazes de distinguir apenas dois níveis de tensão – computadores digitais binários.

Esses sinais elétricos são tensões que assumem dois diferentes valores: um valor positivo (hoje, nos PCs, cerca de +3V) para representar o valor binário 1 e um valor aproximado a 0V para representar o valor binário 0.

Na realidade, esses valores não são absolutos, e sim faixas de valores com uma margem de tolerância (entre + 2,5 e + 3,5 V para 1 e entre 0 e + 0,5 V para 0). A lógica que permite aos computadores operar baseados nesses dois valores é chamada Álgebra de Boole, em homenagem ao matemático inglês George Boole (1815–1864).

- BIT: contração de BInary DigiT e representa um dos valores possíveis em binário, 0 ou 1.
- BYTE: é um grupo de 8 bits (é bom lembrar que $2^3 = 8$). Em um byte, há $2^8 = 256$ combinações, portanto, pode-se representar 256 diferentes valores, desde 00000000 até 11111111.

Na informática, a expressão kilo (abreviada por k) equivale a 2^{10} . Dessa forma, 1 kilobit (1 kb) equivale a 2^{10} bits (1024 bits) e 1 kilobyte (1 kB) equivale a 2^{10} bytes, ou seja, 1024 bytes ou, ainda, 8.192 bits. Da mesma forma, a expressão mega equivale a 2^{20} , ou seja, $2^{10} \times 2^{10} = 1.048.576$. Dessa forma, 1 megabit (1 Mb) equivale a 2^{20} bits, ou seja, 1024 kb ou 1.048.576



bits, e 1 megabyte equivale a 2^{20} bytes, ou seja, 1.048.576 bytes. Seguem-se 1 giga, equivalente a 2^{30} ou 1024 megas, 1 tera, equivalente a 2^{40} ou 1.024 gigas, e 1 peta, equivalente a 2^{50} ou 1.024 teras.

Tipo	Capacidade	Velocidade	Custo	Localização	Volatilidade
Registrador	Bytes	Muito alta	Muito alto	Processador	Volátil
Cache	MB	Alta	Alto		
Memória principal	GB	Média	Médio	Placa-mãe	
Memória secundárias	GB/TB	Baixa	Baixo	Periférica	Não volátil



Unidades de transferência e armazenamento

As unidades de transferência de dados são medidas da seguinte forma:

- Em memórias primárias: é o número de bits que podem ser lidos/escritos de cada vez (palavra).
- Em memórias secundárias: é a quantidade de blocos de dados lidos/escritos por vez.

Um termo comumente empregado é a palavra, que representa a unidade de informação da CPU e da memória principal. O conceito mais usado define palavra como sendo a capacidade de manipulação de bits do núcleo do computador (CPU e MP). Pressupõe-se aqui que todos os elementos do núcleo do computador (que incluem o tamanho da ULA, registradores da CPU e o barramento de dados) tenham a mesma largura (processem simultaneamente o mesmo número de bits), o que nem sempre acontece. Muitas vezes encontram-se computadores cujo tamanho da ULA e dos registradores não é o mesmo dos barramentos.

Dessa forma, encontram-se especificações de computadores de 64 bits, mesmo quando seu barramento de dados é de 32 bits. Concluindo, deve-se analisar caso a caso, porque a simples menção ao tamanho da palavra não é uma terminologia que permita definir de forma conclusiva a arquitetura do computador.

Quanto ao armazenamento, os dados apresentam outra forma. Em geral, o termo "célula" é usado para definir a unidade de armazenamento, e o termo "palavra" para definir a unidade de transferência e processamento, significando na prática quantos bits o computador movimenta e processa em cada operação. Não confundir: célula não é sinônimo de palavra, embora em algumas máquinas a palavra seja igual à célula.

Hierarquia de memórias

As restrições no projeto de memória de um computador podem ser sumarizadas em três questões: Qual sua capacidade de armazenamento? Qual sua velocidade de acesso? Qual seu preço por bit armazenado?



Existe uma variedade de tecnologias usadas para implementar os sistemas de memória. Nesse espectro de tecnologias, podem ser consideradas as seguintes relações:

- Tempo de acesso mais curto, maior o custo por bit.
- Maior capacidade, menor o custo por bit.
- Maior capacidade, maior o tempo de acesso.



Um projetista gostaria de usar as tecnologias de memória que disponibilizam uma elevada capacidade de memória, tanto porque a capacidade é necessária como pelo baixo custo por bit. Contudo, para atingir os requisitos de rendimento, necessita usar memórias dispendiosas, de relativamente baixa capacidade de armazenamento e baixo tempo de acesso.

A solução desse dilema não está dependente de um simples componente de memória ou da tecnologia, mas do emprego de uma hierarquia de memória.

1. Redução do custo/bit
2. Aumento de capacidade
3. Aumento do tempo de acesso
4. Redução da frequência dos acessos à memória pelo processador

Para entender um pouco mais sobre hierarquia de memórias clique no botão abaixo e visualize a imagem interativa.

IMAGEM

INTERATIVA

(http://ead.uninove.br/ead/disciplinas/web/_g/arco80_100/a13i01_arco80_100.htm)

Para que um dado localizado na memória virtual possa ser utilizado no processamento é preciso que passe por cada nível da hierarquia. A razão principal é que o custo por bit de uma tecnologia de memória é geralmente proporcional à sua velocidade. Memórias rápidas, como SRAM, tendem a ter alto custo por bit, tornando proibitivamente caro construir a memória de um computador totalmente com esses dispositivos.

Na hierarquia, os níveis mais próximos ao processador, como a cache, contêm uma quantidade relativamente pequena de memória, que é implementada numa tecnologia de memória rápida, de modo a fornecer um baixo tempo de acesso.

O objetivo de uso da hierarquia de memória é manter os dados que serão mais referenciados por um programa nos níveis superiores, de modo que a maioria das solicitações à memória possa ser tratada no(s) nível(is) superior(es). Por essa razão, o nível mais alto (cache) é geralmente implementado utilizando SRAM, com o emprego de blocos pequenos (normalmente entre 32 e 128 bytes) controlados por hardware. Já a memória principal é geralmente construída com DRAM, com o tamanho de blocos grandes (vários kilobytes) e controle pelo sistema operacional. E, finalmente, a memória virtual é usualmente implementada utilizando discos que contêm todos os dados do sistema de memória.



Quando uma solicitação de memória é enviada para a hierarquia, o nível mais alto é verificado para ver se contém o endereço. Se for assim, a solicitação é completada. Caso contrário, o próximo nível mais baixo é verificado, com o processo sendo repetido até que o dado seja encontrado ou o nível mais baixo da hierarquia seja atingido, no qual se tem a garantia de que o dado está contido.

Com isso, um bloco de posições sequenciais contendo o endereço referido é copiado do primeiro nível que contém aquele endereço para todos os níveis acima dele. As razões para a utilização de blocos são:



- Muitas tecnologias de armazenamento (DRAM modo paginado, discos rígidos) permitem que várias palavras sequenciais de dados sejam lidas e escritas em menos tempo que um número igual de palavras localizadas aleatoriamente, tornando mais rápido transferir um bloco de vários bytes de dados que buscar cada byte individualmente.
- A maioria dos programas apresenta localidade de referência: as referências à memória que ocorrem próximas no tempo tendem a ter endereços que são próximos uns aos outros, fazendo com que seja provável que outros endereços dentro do bloco sejam referenciados em breve, após um primeiro acesso a um endereço no bloco.

Referências

STALLINGS, Willian. Arquitetura e organização de computadores. 5. ed. Prentice Hall. São Paulo, 2006.

TANENBAUM. Andrew S. Organização estruturada de computadores. 5. ed. Rio de Janeiro: LTC, 2007.

MACHADO, Francis B.; MAIA, Luiz P. Arquitetura de sistemas operacionais. 4. ed. Rio de Janeiro: LTC, 2007.

WEBER, Raul Fernando. Arquitetura de computadores pessoais. 2. ed. Porto Alegre: Sagra Luzzatto, 2003.

_____. Fundamentos de arquitetura de computadores. 3. ed. Porto Alegre: Sagra Luzzatto, 2004.



Avalie este tópico



ANTERIOR

<

Estrutura básica do computador; busca e execução de instruções; interrupções; barramentos

Biblioteca

(https://www.uninove.br/conf/indice)

uninove/biblioteca/sobre-a-biblioteca/apresentacao/)

Índice

≡

Memória principal de semicondutor - memória cachê - memórias magnéticas e memórias ópticas

Ajuda?

(https://ava.uninove.br/curso=)

>

Portal Uninove
(<http://www.uninove.br>)
[Mapa do Site](#)

