

Memória

Em um sistema de computação temos a UCP se comunicando com a memória e os dispositivos de E/S.

Podemos fazer um paralelo do acesso à memória com um carteiro entregando cartas ou um acesso à biblioteca.

Duas ações básicas podem ser executadas:

- armazenamento (escrita ou gravação (write))
- recuperação (leitura (read))

Em sistemas de computação a unidade básica é o bit que pode ter 2 valores (0 ou 1). Mas precisamos de mais valores para representar informação: números, letras do alfabeto, símbolos.

Para isso, os bits são agrupados em grupos de bits que são armazenados na memória e são acessados sempre em grupo.

Para acessar um grupo de bits, utiliza-se um endereço que varia de 0 a (N-1) onde N é o número total de grupos existentes.

Exemplo: Ler o endereço 0

Escrever no endereço 4 o grupo 0101

```
end 0 |-----|
      | M bits |
end 1 |-----|
      | M bits |
end 2 |-----|
      | M bits |
      |-----|
      | .       |
      |-----|
end N-1 |bm-1|bm-2|...|b0|
        |-----|
```

Operação de escrita

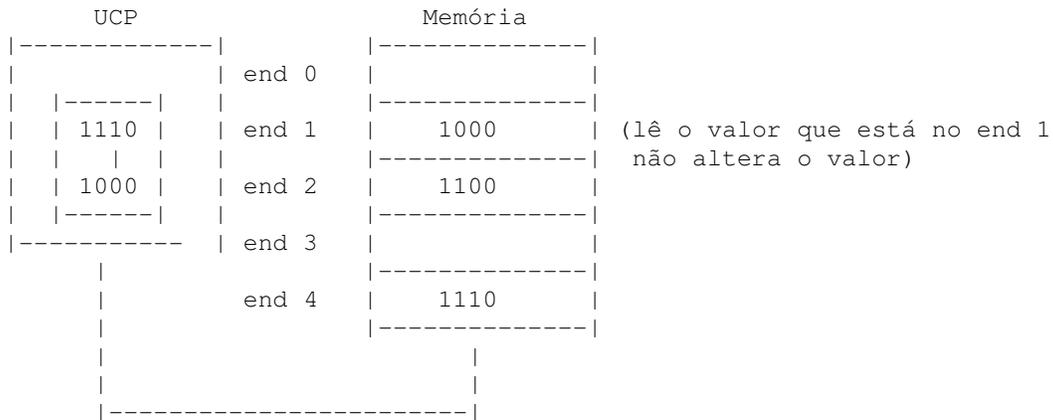
Escreve no endereço 4

UCP		Memória	
-----	end 0	-----	
-----	end 1	-----	
1110	end 2	1000	
-----	end 3	-----	
	end 4	1100	

		0001 -> 1110	(substitui o valor antigo pelo novo valor)

Operação de leitura

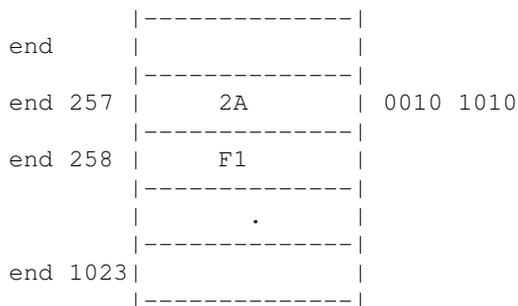
Lê do endereço 1



Memória

- Dados e instruções são armazenados na memória, lidos e escritos pela UCP. É a UCP que manipula a memória, lendo ou escrevendo dados. Os dados trafegam no barramento de dados.
- Instruções são armazenadas de forma seqüencial (mas podem não ser executadas da mesma forma).
- Memória organizada em grupos de bits chamados células, onde cada célula pode ter de M bits.
- Cada célula (ou grupo de bits) possui um endereço associado.
- O grupo de bits transferido na leitura e na escrita entre a memória e a UCP se chama palavra. O barramento de dados deve ser capaz de transportar uma palavra.
- O acesso (leitura e escrita) não depende do tamanho da célula mas do tamanho da palavra. O tamanho da palavra é a quantidade de bits transferida na leitura e na escrita.
- Organização utilizada atualmente: cada célula possui 8 bits e o tamanho da palavra varia entre 16, 32 e 64 bits, isto é, é possível ler mais de uma célula por vez.

Exemplo de organização de memória: Memória com 1Kbytes= 2^{10} =1024 bytes, cada célula possui 8 bits.

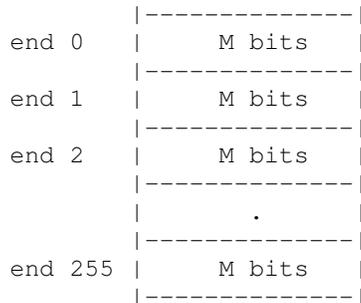


Organização da Memória

Parâmetros a serem definidos na organização da memória

- Quantidade de bits de uma célula: geralmente utiliza-se um byte = 8 bits

Exemplo de organização de memória: Memória com 256 células = 2^8
e cada célula pode ter $M = 12, 16$ ou 8 bits (tamanho da célula)



- Quantidade de células X endereçamento
Endereço é utilizado para endereçar uma célula e é formado por um grupo de bits.
Podemos ter máquinas com mesmo tamanho de célula e quantidade diferente de células.

Exemplo 1:

Memória com 8K células de 8 bits cada uma.

Qual a capacidade da memória? $2^3K \times 8$ bits

$$2^{13} \times 2^3 \text{ bits} = 2^{16} \text{ bits} = 64 \text{ K bits}$$

$$(\text{se utilizarmos bytes como medida de capacidade}) = 2^6 \text{ K} / 2^3 = 2^3 \text{ K bytes}$$

Exemplo 2:

Memória com 128K células = 2^7 células de 8 bits cada uma.

Qual a capacidade da memória? $128K \times 1 = 128K$ bytes

Capacidade da memória $T = N \times M$, onde N é o número de células e M o número de bits de cada célula.

Exemplo: $N=1024$ e $M=4$ bits

$$T=1024 \times 4 = 4096 \text{ bits} = 2^{10} \times 4 \text{ bits} = 4K \text{ bits} = 512 \text{ bytes} \quad (4096 / 8)$$

$$1K=2^{10}=1024, \quad 1M=2^{20}=1.048.576, \quad 1G=2^{30}=1.073.741.824$$

Exemplo: Se cada célula da memória pode armazenar 8 bits e um int em C possui 4 bytes (i.e., 4×8 bits), então cada acesso à memória traz 4 células para a UCP (são necessárias 4 células para armazenar um inteiro).

Se M (tamanho da célula) = 4 bits, quantos valores diferentes podemos ter armazenados? 2^4 valores

Tamanho do endereço:

Cada endereço da memória é obtido por E bits que representa o tamanho do endereço. Então, para armazenar N endereços no REM (registrador de endereços de memória) gastamos E bits (ou linhas de endereço), logo $N=2^E$.

Exemplo: Se o número total de células é $N=1024$, quantos bits usamos para endereçamento?

$1024 = 2^{10}$, então $E=10$. São necessários 10 bits para armazenar (ou linhas de endereço)

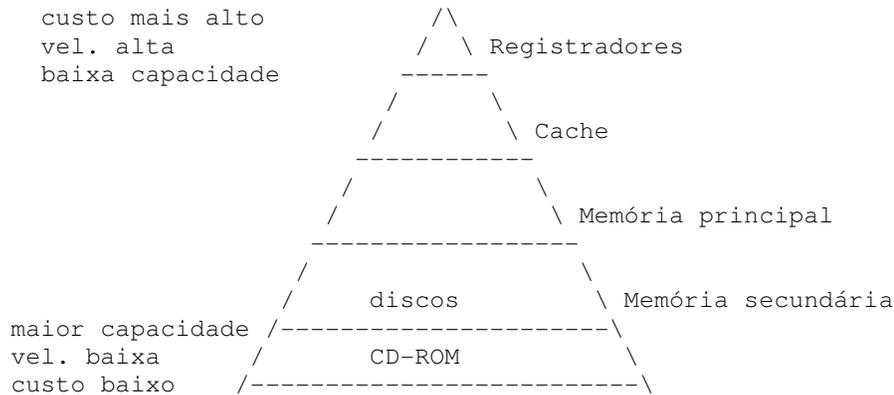
Exemplos:

1. Espaço de endereçamento = 8K células e cada célula pode armazenar 8 bits. Qual a capacidade da Memória e quantas linhas de endereço devem ser utilizadas ? $N = 8K = 2^3 \times 2^{10} = 2^{13}$ e $M = 2^3$
(capacidade) $T = N \times M = 2^{13} \times 2^3 = 2^{16} = 64K$ bits
(linhas de endereço) $N = 2^{13}$, logo teremos que ter 13 linhas para endereço
2. Uma memória pode armazenar no máximo 512K bits e cada célula possui 8 bits. Qual é o número máximo de células possíveis e quantos bits devem ser utilizados para endereçar este número de células?
(células) $T = N \times M \Rightarrow N = T/M = 512 \times 2^{10} / 8 = 2^9 \times 2^{10} / 2^3 = 2^{16} = 2^6 \times 2^{10} = 64K$ células.
(Bits para endereçamento) $N = 64K = 2^{16}$, logo precisaremos de 16 bits
3. Um computador possui capacidade máxima de armazenamento de 512K células, cada uma de 16 bits.
 - a) Qual o número mínimo de bits que o barramento de endereços deve ter?
Se temos 512 K células $= 2^9 \times 2^{10} = 2^{19}$, teremos 19 bits para o barramento de endereços.
 - b) E qual é o número mínimo de bits que o barramento de dados deve ter, supondo que duas células são acessadas em cada operação de leitura/ou escrita? Se em cada acesso à memória, duas células são transferidas, então o barramento de dados deve ter o tamanho igual a 2 x o tamanho de cada célula: (2×16) 32 bits.
 - c) Qual o maior endereço da Memória e qual a capacidade da Memória?
O maior endereço da Memória é $N-1=2^{19}-1=524.287$
Capacidade da Memória: $T = N \times M = 2^{19} \times 16 = 2^{23}$ bits $= 2^{20}$ bytes $= 1M$ bytes
4. Suponha que uma máquina possui um barramento de dados de 64 bits e um barramento de endereços de 24 bits. Sabe-se que em cada acesso são lidas duas células da memória.
 1. Qual a capacidade máxima de endereçamento da máquina?
Se temos 24 bits para endereço, podemos endereçar 2^{24} células $= 16M$ células ou 16M endereços
 2. Qual o tamanho de célula da máquina? Em cada acesso são lidas 2 células, e como o barramento de dados é de 64 bits, temos que cada célula possui 32 bits.
 3. Qual a capacidade máxima de armazenamento da memória em bits ?
Capacidade = $32 \times 16M = 512$ Mbits

Hierarquia de memória

Vários tipos de memória:

- Mais rápida, mais cara e menor capacidade
- Mais lenta, mais barata e maior capacidade



Parâmetros de memória

- Tempo de acesso: tempo decorrido entre o pedido de dado à memória (endereço) e a disponibilização do dado pela memória.
Ex: RAM (Random Access Memory) – neste tipo de memória o tempo independe da posição do dado, então é possível escolher o endereço aleatoriamente (random)
Disco e CD-ROM – o tempo depende da posição do dado então o endereço é seqüencial.
- Capacidade: medida em bytes (8 bits)
 $1K=2^{10}$, $1M=2^{20}$, $1G=2^{30}$, $1T=2^{40}$
- **Volatilidade:** diz respeito à capacidade de manter dados sem energia elétrica
 - Memória volátil: não consegue manter os dados sem energia elétrica
- Tecnologia: semicondutores, magnéticos, óticos

Registradores

Unidade de memória que fica dentro da UCP e que armazena dados e instruções.

Exemplos: registrador de instrução, registrador de dados

- Tempo de acesso: tecnologia igual a UCP, muito rápido. Um ciclo de máquina, 1GHZ, 1ns
- Capacidade: 1 único dado ou instrução, endereço, por registrador (pouca capacidade.)
Exemplo: Intel: 10 registradores de 32 bits
MIPS (Silicon Graphics): 32 registradores de 32 bits
- Voláteis e caros

Memória principal

Memória utilizada para armazenar o programa a ser executado. A UCP acessa a memória para ler as instruções que serão executadas e os dados que irá manipular.

- Tempo de acesso: grande, pode variar de 50 a 150 ns.
- Capacidade: bem grande: 64 MB, 128 MB, 256 MB, 1GB
- Voláteis e baratas
- Tecnologia DRAM (Dynamic RAM)

RAM dinâmica x RAM estática: RAM estática a informação permanece armazenada enquanto tiver informação e a RAM dinâmica a informação precisa ser restabelecida de tempos em tempos.

Memória Cache

O ciclo de instrução é composto do ciclo de memória mais o ciclo de execução, isto é, é necessário buscar os dados na memória e depois executar a instrução.

A tecnologia possibilitou um aumento de velocidade muito maior para UCP (onde os dados são executados) do que para memória principal (de onde os dados são lidos), causando um gargalo.

Solução: colocar uma memória intermediária conhecida como cache. A cachê tem uma velocidade maior que a memória principal porém é mais cara.

A cachê pode estar tipicamente em dois níveis:

Nível L1: dentro do processador.

Nível L2: fora do processador na placa mãe

- Tempo de acesso: menor que memória principal, maior que registradores, 10ns.
- Capacidade: se muito grande, muito cara. Se muito pouca, baixo desempenho pois tem que ficar buscando na memória principal toda a informação.
Valores típicos: 16K para L1 e 64K a 2MB para L2.
- Voláteis e caras
- Tecnologia SRAM (Static RAM)

Memória secundária

Memória que permite o armazenamento permanente de dados e instruções.

Pode ser constituída de vários dispositivos: discos rígidos, disquetes, fitas, CD-ROMs.

- Tempo de acesso: dispositivos eletromecânicos possuem tempo maior que dispositivos semicondutores.
Discos rígidos: 8 a 15ms
CD-ROMs: 120 a 300ms
Fitas: ordem de segundos
- Capacidade: bem maior que a memória principal: 100GBytes, 1 Terabyte
- NÃO Voláteis e mais baratas
- Magnético e ótico