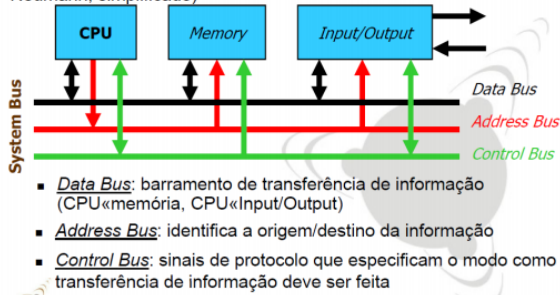


Processadores

Modelo de Von Neumann

- Arquitectura básica de um sistema computacional (modelo de von Neumann, simplificado)



Passos para a CPU executar cada instrução:

- Acede à próxima instrução existente em memória e coloca-a no registo de instrução;
 - Atualiza o contador de programa (instructor pointer) para que ele aponte para a instrução seguinte;
 - Determina o tipo de instrução (se a instrução utiliza dados de memória, determina onde estão);
 - Acede aos dados, se os houver, colocando-os em registos internos da CPU;
 - Executa a instrução;
 - Armazena os resultados nos locais apropriados;
- * A cada impulso de relógio a CPU pode desencadear um novo ciclo de aquisição e execução de instruções;
- * O desempenho da CPU é medido em MIPS (Milhões de instruções por segundo);

Aspetos a considerar na Evolução dos CPU:

1946 – 1ª Geração – Válvulas e Cartões perfurados;
1958 – 2ª Geração – Transístor e linguagens de alto nível, primeiros computadores comerciais;
1965 – 3ª Geração – Circuitos Integrados e Minicomputadores
1971 – Gerações Seguintes – Microprocessadores

4004: Primeiro verdadeiro microprocessador (ainda em 1971)

registos: 4 bits

bus dados: 4 bits

bus endereços: 12 bits

+ ou – 2300 transístores -> Lei de Moore : o número de transístores num circuito integrado duplica a cada 2 anos

8008

registos: 8 bits

bus de dados: 8 bits

bus de endereços: 14 bits

8080

registos: 16 bits

- estrutura interna que garantia uma compatibilidade descendente (Explosão dos PC's!)

8086

registos: 16 bits

bus de dados: 16 bits

bus de endereços: 20 bits

- memória concebido como superconjunto -> segmentos de 64 KB

8088 – Versão low cost do 8086

80286

- permitia o uso de memória virtual – tecnicamente permite a execução de processos em que parte se encontra em memória principal e outra parte em disco;
- possibilidade de controlar até 1GB de memória total – 16MB físicos e 1088MB virtuais

Memória virtual – permite a execução de processos em que parte se encontra em memória principal e outra parte em disco, permite o uso mais eficiente da memória física;

Memória física – RAM – blocos de tamanho fixo, molduras de páginas;

Memória lógica – disco – dividida por blocos do mesmo tamanho, páginas;

MMU (Memory Management Unit) - tem a responsabilidade de traduzir endereços lógicos em endereços físico.

Para manter a compatibilidade com os chips antigos este tem 2 modos operacionais:

Modo Real: reproduz o esquema de operação do 8086; herda as limitações do 8086 como a barreira de 1MB de memória;

Modo Protegido: tirava partido no novo esquema de memória, embora não compatível com os programas anteriores para 8086, permitia o uso dos 16MB de memória real e 1GB de memória virtual;

80386 DX

registos: 32 bits

bus dados: 32 bits

bus endereços: 32 bits (4GB memória física e 64TB memória virtual)

bus dados = bus endereços → permite abandonar artifícios de aritmética booleana para calcular os endereços

- a divisão de memória em segmentos é possível e estes não estão restritos aos 65KB habituais podendo ter qualquer tamanho desde que inferior a 4GB;

- incorpora 16 bytes de Pre-fetch Cache memory (armazena as próximas instruções do programa) ;

- inicia em modo real , podendo trocar de modo sem ser reiniciado;

Novo Modo : Virtual – possibilita a simulação de vários 8086 (multitarefa)

80 386 SX – mais barato – 16 bits em vez dos 32 de bus dados

80 386 SL – destinado a portáteis - opera a uma voltagem mais baixa, consomem menos energia

80 386 SLC – tem cache 8 KB

80 486 DX – encapsulamento PGA com processador matemático integrado (FPU) ; cache interna de 8KB , utilização do recurso “ multiplicação do relógio “ ; necessidades de coolers .

80 486 SX – sem FPU

80 486 SL – menos utilização de energia ; gestão energia – desativa periféricos quando não são necessários

80 486 SLC2 – mais barato, tecnologia de 0.7/ μm (em vez de 1 μm)

Pentium (1ª Geração) – bus dados de 64 bits ; cache interna (L1) 16KB ; cache externa (L2) entre 128KB e 2MB

→ **Pipelining** – Executar várias fases simultaneamente. Fases seguintes não dependem das anteriores e o tempo destas deve ser equilibrada. Limitações: Saltos condicionais atrasam o sistema e atribuições sucessivas atrasam o sistema;

→ **Super Pipelining** – aproveitar o bordo descendente do relógio, podia executar 2 instruções por ciclo de relógio (em certas condições);

→ **Super Escalar** – o processador reordena o código para evitar situações de dependência sucessiva; Necessários múltiplas capacidades na CPU; Limitações:

True Data Dependency – Tem de se esperar o resultado de uma operação para se poder executar a outra;

Procedural Dependency – Saltos Condicionais

Pentium (2ª Geração) – tipo de construção de portas lógicas BiCMOS (muita alta velocidade de operação) ; combinam 2 tipo de transístores:

CMOS: mais baratos , mais densidade de encapsulamento, menos consumo de potência estática

Bipolar: mais velocidade de comutação , menos densidade de encapsulamento, mais dissipação de potencia

Pentium Pro (3ª Geração)

- Arquitetura Híbrida (tecnologia abandonada mais tarde) : núcleo RISC (Reduced Instruction Set Computer) , decodificador CISC (Complex Instruction Set Computer - dimensão da instrução)
- Execução dinâmica que visa otimizar o pipelining → execução especulativa (pipelining dinâmico)
- 3 instruções por ciclo de relógio

Pentium MMX – otimizado para acelerar aplicações multimédia a preços baixos

Pentium II

- Encapsulamento SECC
- Registos de 64 bits
- Bus endereçamento 36 bits
- 5 instruções em simultâneo
- Não é mais que um Pentium Pro com acrescento das instruções MMX , permite juntar , de forma económica, memoria cache L2 ao processador .

Pentium III

- 70 Novas instruções multimédia SSE (Streaming Simd Extensions)
- DIB (Dual Independent Bus) – acesso simultâneo á memória principal (FSB) e á cache L2 (BSB)

Pentium III XEON – Destinado a servers (múltiplos processadores)

Pentium IV

- FSB – elevada velocidade
- Hyperpipeline – número de estágios do pipeline duplicado
- Execução dinâmica avançada – até 126 instruções em vez das 42 do Pentium III
- Motor de execução rápida - ALU's executam até 2 operações por ciclo em determinadas instruções
- Execution trace cache – armazena as instruções já decodificadas

Hyper – Threading : Divisão de funções de um programa por vários núcleos dentro do mesmo processador. Em termos de arquitetura um processador com Hyper-threading consiste em dois processadores lógicos cada um com o seu estado. Estes processadores lógicos partilham os recursos de execução.

Pentium D – Dual Core, multiprocessamento real e memoria cache partilhada

Pentium Core 2 Quad – 4 nucleos de execução, 2 x 4 MB cache L2 (cada um dos 2 tem 4MB ao seu dispor)

Core i3 – controladora gráfica PCI-Express integrada, mais autonomia

Core i7 / Core i9 – Quick Path – Substituição do FSB, que em arquiteturas com mais núcleos do processador prejudicava o desempenho do sistema, por um novo barramento de links independentes.

Memórias Tipos:

Função : Apenas de Leitura / Leitura e Escrita

Acesso: Acesso Aleatório / Acesso Sequencial

Manutenção dos Dados: Voláteis / Não Voláteis

ROM – só de leitura , não volátil

PROM – programável , mecanismo de queima de fusíveis

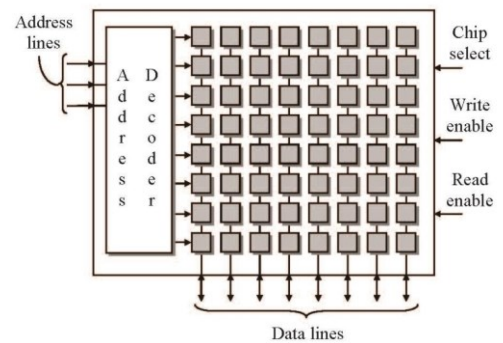
EPROM – apagável (erasable) , através de luz ultravioleta

EEPROM – electronic EPROM

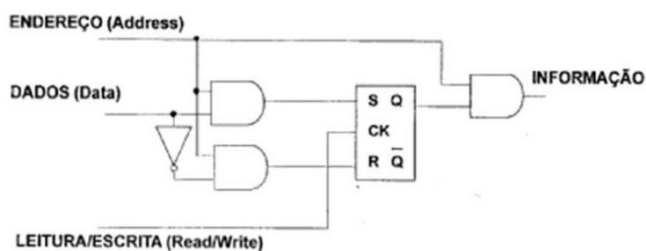
RAM – volátil , acesso aleatório

SRAM – RAM estática

Organização da memória



Exemplo de uma SRAM 1 bit



O flip-flop é o elemento básico de uma SRAM

Decodificador de Endereços

Memória com 4 localizações

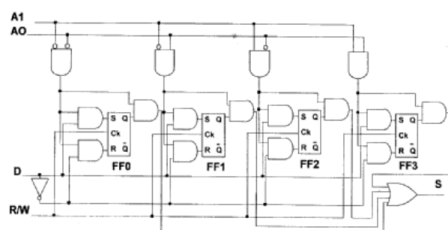


Fig. 1.7 — Memória SRAM constituída por quatro elementos.

Legenda

- A_1A_0 : Linhas de endereço (Address)
- D: Dado (1 bit)
- R/W: Linha de leitura/escrita

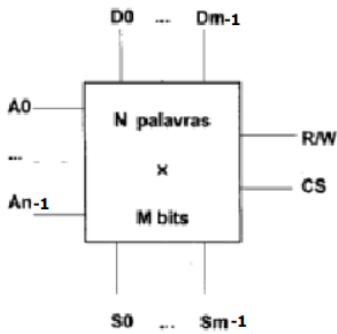
- A seleção de cada elemento é feita por uma malha de portas AND.

Arquitetura unidimensional de seleção de endereço - Complexidade não está no número de AND mas sim no número de entradas deles;

(alternativa)

Arquitetura Bidimensional - O endereço (se par) é cortado ao meio , o numero de AND aumenta mas o numero de entradas é sempre 2

Representação simbólica de um circuito integrado RAM



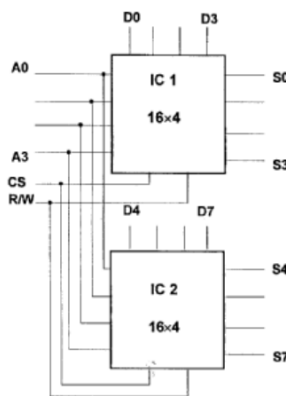
Ligação de memórias em paralelo – usando o CS (Chip Select) é possível:

- aumentar o número de palavras (células, linhas) : o CS liga as 2 memórias , saídas e dados não são comuns;
- aumentar o comprimento (número de bits, colunas) : o CS ativa uma memória de casa vez, as saídas e dados são comuns;
- aumentar ambos

Exemplos

• Exemplo1

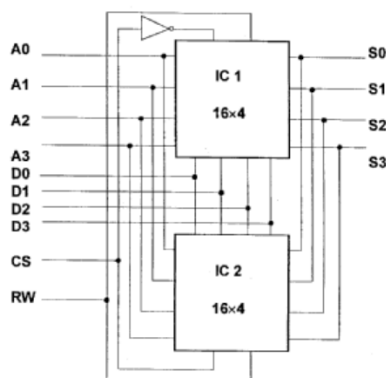
- $A_3A_2A_1A_0$: Linhas de endereço (*Address*)
- R/W: Linha de leitura/escrita
- $D_7D_6D_5D_4D_3D_2D_1D_0$: Dados (entrada)
- $S_7S_6S_5S_4S_3S_2S_1S_0$: Dados (saída)



exemplo1: mantendo o mesmo número de localizações ou palavras, duplica-se o número de bits por palavra, ou seja , o seu comprimento; o CS ativa as 2 memórias simultaneamente, as saídas e os Dados não são comuns.

• Exemplo2

- $A_3A_2A_1A_0$: Linhas de endereço (*Address*)
- R/W: Linha de leitura/escrita
- $D_3D_2D_1D_0$: Dados (entrada)
- $S_3S_2S_1S_0$: Dados (saída)



exemplo2: mantendo o mesmo comprimento das palavras aumenta-se o seu número; o CS ativa 1 memória de cada vez ; as saídas e os dados são comuns – barramento comum: multiplexagem.

DRAM - RAM dinâmica

- o condensador é o elemento básico de uma DRAM;
- Mais lenta , menos complexa
- Solução menos eletrônica, mais elétrica

Problema: Não são isolantes perfeitos e só aguentam a carga por uns milissegundos

Solução: Mecanismos de refrescamento de memória (processo destrutivo) → *daqui é que resulta o nome de memórias dinâmicas!*

Vantagens da DRAM sobre SRAM:

+ barata , + densidade de integração, - consumo, + dissipação eficiente

Desvantagens da DRAM sobre SRAM:

+lentas, +complexas

Solução : Modelos Híbridos!

SRAM → cache

DRAM → memórias principais

Medidas de Velocidade da memória

Tempo de acesso

Velocidade do Bus

Largura do Bus

Latência RAS (Row Access Strobe)/CAS (Column Access Strobe)

Memórias Regulares – acesso feito enviando 1º o RAS e depois o CAS

Assíncronas:

FPM (Fast Page Mode) – Ativação RAS e depois múltiplos endereços CAS (não é verdadeiramente aleatório) .

EDO (Extended Data Output) – O mesmo que a FPM no entanto um acesso aos dados pode ser iniciado antes que o anterior termine.

BEDO (Burst EDO) – Bursting / Pipelining – em vez de aceder a 1 bit apanha um bloco de info com endereços sucessivos

SDRAM (Synchronous Dynamic RAM) – Utiliza um sinal de relógio adicional para os sinais de controlo

- Interleaving – acesso a dois ou mais bancos de memória a fim de encurtar os ciclos de memória

- Bursting

DDR SDRAM (Double Data Rate Synchronous Dynamic RAM) – transações em ambos os bordos do relógio

DDR II SDRAM

- menor consumo de energia

- menor custo de provação