

Arquitetura e Organização de Computadores

Capítulo 2

Evolução e desempenho do computador

slide 1

© 2010 Pearson Prentice Hall. Todos os direitos reservados.

ENIAC – histórico

- Electronic Numerical Integrator And Computer.
- Eckert (aluno) e Mauchly (Professor).
- Universidade da Pensilvânia.
- Tabelas de trajetória para armas.
- Iniciou em 1943.
- Concluído em 1946.
 - Muito tarde para o esforço de guerra.
- Usado até 1955 pelo Exército americano.

ENIAC – detalhes

- Decimal (não binário).
- 20 acumuladores de 10 dígitos.
- Programado manualmente por chaves.
- 18 000 válvulas.
- 30 toneladas.
- 1 500 pés quadrados.
- 140 kW de consumo de potência.
- 5 000 adições por segundo.

Proposta de John von Neumann (em 1945)

- 1. Unidade Lógica e Aritmética (CA)
 - Se um computador deve realizar operações matemáticas, precisa de uma unidade específica
- 2. Memória (M)
 - Se um computador vai realizar operações complexas, precisa de memória
- 3. Unidade de Controle (CC)
 - A sequenciação apropriada de operações será mais eficiente se realizada por uma unidade de controle central

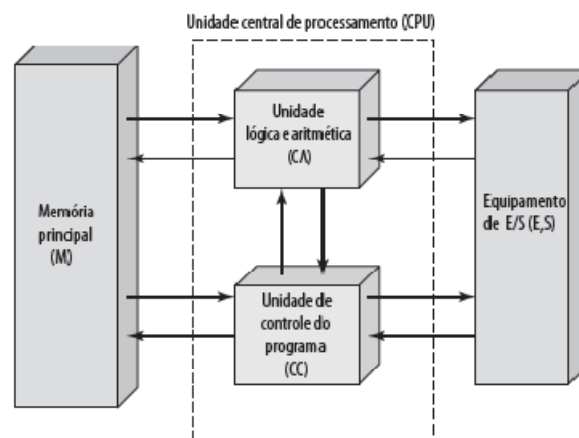
=> Essas 3 partes são os neurônios *associativos* !

Proposta de John von Neumann

- 4. Unidade de Entrada (I)
 - Se um computador deve receber dados do exterior, precisa transferir esses dados através de uma unidade específica
 - 5. Unidade de Saída (O)
 - Se um computador deve retornar resultados para o exterior, precisa de uma unidade específica
- => Essas 2 partes são os neurônios *sensoriais* (E) e *motores* (S)!

CA – Central Arithmetic; M – Memory; CC - Central Control;
I – Input; O – Output.

Estrutura da máquina de von Neumann



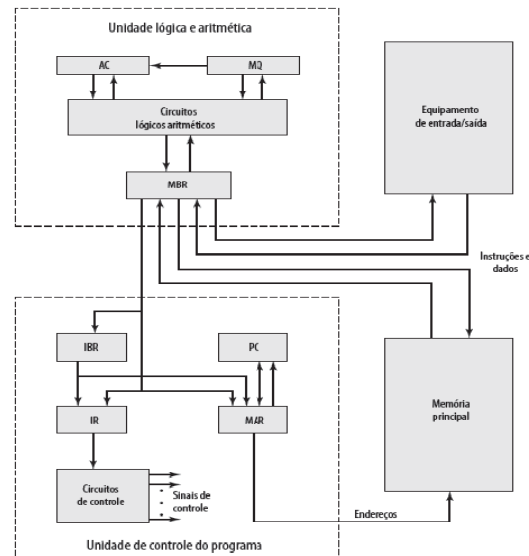
von Neumann/Turing

- Conceito de programa armazenado.
- Memória principal armazenando programas e dados.
- ALU operando sobre dados binários.
- Unidade de controle interpretando e executando instruções da memória.
- Equipamento de entrada e saída operado por unidade de controle.
- *Princeton Institute for Advanced Studies.*
 - IAS
- Concluído em 1952.

IAS – detalhes

- 1000 “palavras” de 40 bits.
 - Número binário.
 - 2 instruções de 20 bits (esquerda e direita)
- Conjunto de registradores (armazenamento em CPU).
 - Registrador de buffer de memória (MBR)
 - Registrador de endereço de memória (MAR)
 - Registrador de instrução (IR)
 - Registrador de buffer de instrução (IBR)
 - Contador de programa (PC)
 - Acumulador (AC)
 - Quociente multiplicador (MQ)

Estrutura do IAS – detalhe

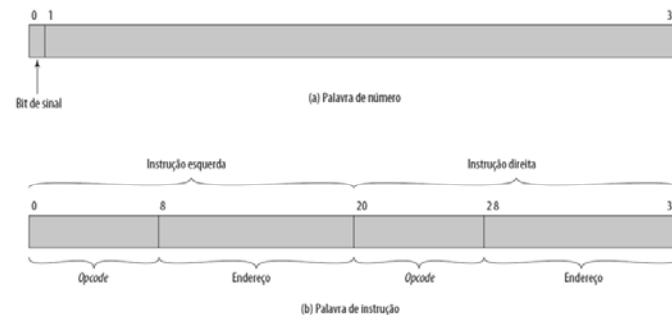


IAS – detalhes

- MBR – contém uma palavra a ser armazenada na memória ou enviada à unidade de E/S, ou é usado para receber uma palavra da memória ou de uma unidade de E/S
- MAR – armazena o endereço na memória da palavra a ser escrita ou lida no MBR
- IBR – registra temporariamente a próxima instrução a ser executada
- IR – armazena o *opcode* da instrução que está sendo executada
- PC – Armazena o próximo par de instruções a ser buscado na memória
- AC e MQ – armazenam temporariamente operandos e resultados de operações da ALU

Formatos de memória do IAS

Figura 2.2 Formatos de memória do IAS



Conjunto de 21 instruções no IAS

- Grupo de Transferência de Dados (7)
 - Ex: LOAD MQ ; transfere MQ para AC
- Grupo de Desvio Incondicional (2)
 - Ex: JUMP M(X,0:19) ; busca metade esquerda de M(X)
- Grupo de Desvio Condicional (2)
 - Ex: JUMP+M(X,0:19) ; busca metade esquerda de M(X), se AC não negativo
- Grupo de operações aritméticas (8)
 - Ex: ADD M(X) ; $AC \leftarrow AC + M(X)$
- Grupo de endereço (2)
 - Ex: STOR M(X,8:19) ; substitui campo de end da esquerda em M(X) por 12 bits à direita de AC

Computadores comerciais

- 1947 – Eckert-Mauchly Computer Corporation.
- UNIVAC I (Universal Automatic Computer).
- Usado pelo Birô do Censo dos EUA para cálculos de 1950.
- Tornou-se parte da Sperry-Rand Corporation.
- Final da década de 1950 – UNIVAC II.
 - Mais rápido.
 - Mais memória.

IBM

- Equipamento de processamento de cartão perfurado.
- Em 1953 – o 701
 - Primeiro computador de programa armazenado da IBM.
 - Cálculos científicos.
- Em 1955 – o 702
 - Aplicações comerciais.
- Levou à série 700/7000.

Uso de Transistores (define 2ª Geração)

- Substituíram as válvulas.
- Menores.
- Mais baratos.
- Menos dissipação de calor.
- Dispositivo de estado sólido.
- Feito de silício (areia).
- Inventado em 1947 na Bell Labs.
- William Shockley e outros.

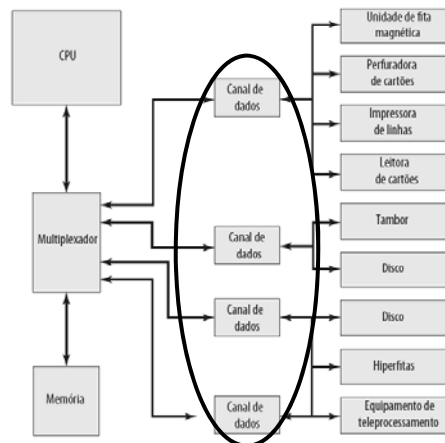
Computadores baseados em transistor

- Máquinas de segunda geração.
- NCR & RCA produziram máquinas com transistor pequeno.
- IBM 7000.
- DEC – 1957 produziu o PDP-1.

Configuração de um IBM 7094

Figura 2.5 Configuração de um IBM 7094

Canais de dados !

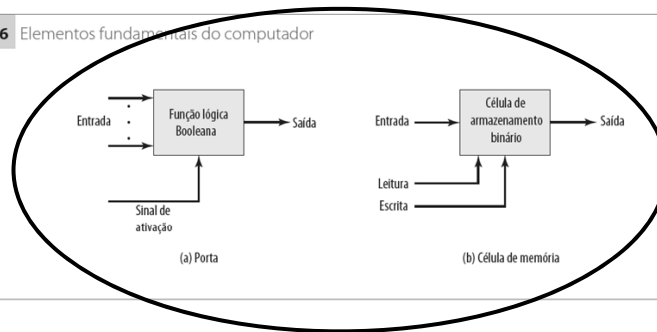


Microeletrônica (define a 3ª geração)

- Literalmente – “pequena eletrônica”.
- Um computador é composto de **portas lógicas, células de memória e interconexões**.
- Estas podem ser fabricadas em um semicondutor.
- Por exemplo, *wafer* de silício.

Elementos Fundamentais do Computador

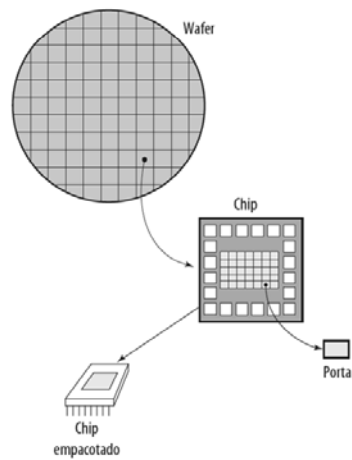
Figura 2.6 Elementos fundamentais do computador



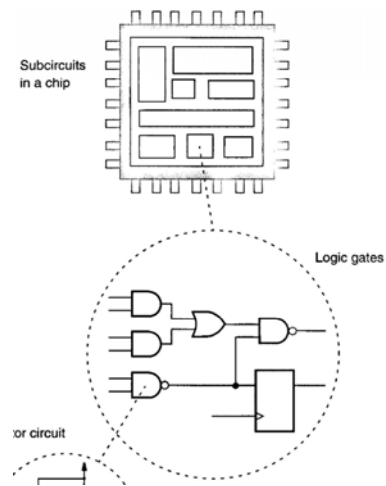
Curso de Eletrônica Digital !

Conceitos de um Circuito Integrado

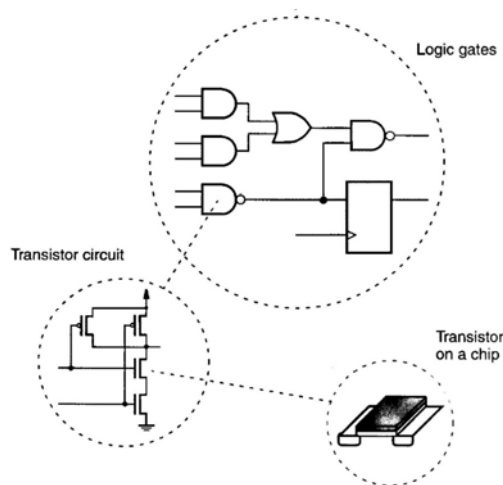
Figura 2.7 Relacionamento entre wafer, chip e porta



Conceitos de um Circuito Integrado



Conceitos de um Circuito Integrado



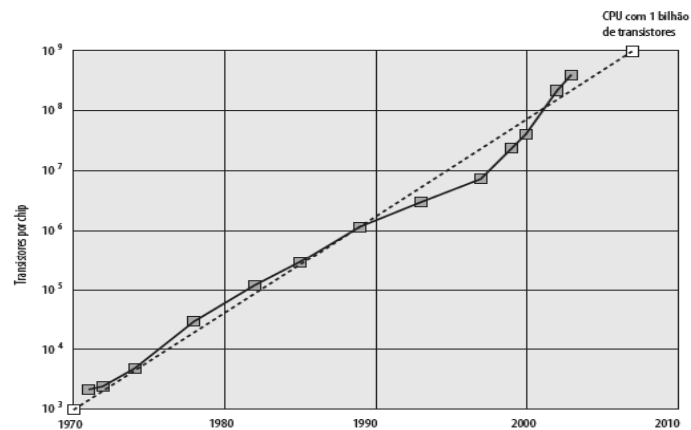
Gerações de computadores

- Válvula – 1946-1957.
- Transistor – 1958-1964.
- Integração em pequena escala (SSI) – 1965 em diante.
 - Até 100 dispositivos em um chip.
- Integração em média escala (MSI) – 1971.
 - 100-3.000 dispositivos em um chip.
- Integração em grande escala (LSI) – 1971-1977.
 - 3.000 – 100.000 dispositivos em um chip.
- Integração em escala muito grande (VLSI) – 1978 -1991.
 - 100.000 – 100.000.000 dispositivos em um chip.
- Integração em escala ultragrande (UVLSI) – 1991.
 - Mais de 100.000.000 dispositivos em um chip.

Lei de Moore

- **“Número de transistores em um chip dobrará a cada ano”**
 - Número de transistores dobra a cada 18 meses.
- Gordon Moore – cofundador da Intel.
- Maior densidade de componentes no chip.
- Desde 1970, isso diminuiu um pouco.
- Custo de um chip permaneceu quase inalterado.
- Maior densidade de empacotamento significa caminhos elétricos mais curtos, gerando maior desempenho.
- Menor tamanho oferece maior flexibilidade.
- Redução nos requisitos de potência e resfriamento.
- Menos interconexões aumenta a confiabilidade.

Crescimento na contagem de transistores da CPU



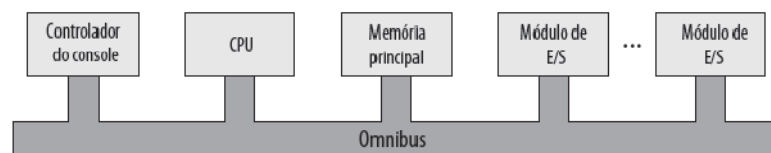
IBM série 360

- 1964.
- Substituiu (incompatível com) série 7000.
- Primeira "família" planejada de computadores.
 - Conjuntos de instruções semelhantes ou iguais.
 - SO semelhante ou igual.
 - Velocidade aumentada.
 - Número cada vez maior de portas de E/S (ou seja, mais terminais).
 - Tamanho de memória crescente.
 - Maior custo.
- Estrutura comutada multiplexada.

DEC PDP-8

- 1964.
- Primeiro minicomputador (nome deve-se à minissaia!).
- Não precisava de sala resfriada.
- Pequeno o suficiente para ser colocado sobre uma bancada de laboratório.
- US\$ 16,000.00.
 - Centenas de milhares de US\$ do IBM 360.
- Aplicações embutidas & OEM.
- Estrutura de barramento.

Estrutura de barramento do DEC PDP-8



Diferente da estrutura comutada central da IBM !

Memória semicondutora

- 1970.
- Fairchild.
- Tamanho de um único core (anel ferromagnético).
 - Ou seja, 1 bit de armazenamento do core magnético.
- Chip de memória semicondutora mantinha 256 bits.
- Leitura não destrutiva.
- Muito mais rápida que o core.
- Capacidade dobra aproximadamente a cada ano.

Intel (microprocessadores)

- 1971 - 4004
 - Primeiro microprocessador.
 - Todos os componentes da CPU em um único chip.
 - 4 bits.
- Acompanhado em 1972 pelo 8008.
 - 8 bits.
 - Ambos projetados para aplicações específicas.
- 1974 - 8080.
 - Primeiro microprocessador de uso geral da Intel.

Evolução dos microprocessadores da Intel

Tabela 2.6 Evolução dos microprocessadores Intel

(a) Processadores da década de 1970

	4004	8008	8080	8086	8088
Introduzido	1971	1972	1974	1978	1979
Velocidades de clock	108 kHz	108 kHz	2 MHz	5 MHz, 8 MHz, 10 MHz	5 MHz, 8 MHz
Largura do barramento	4 bits	8 bits	8 bits	16 bits	8 bits
Número de transistores	2 300	3 500	6 000	29 000	29 000
Dimensão mínima da tecnologia de fabricação (μm)	10		6	3	6
Memória endereçável	640 bytes	16 KB	64 KB	1 MB	1 MB

Evolução dos microprocessadores da Intel

(continuação)

(b) Processadores da década de 1980

	80286	386TM DX	386TM SX	486TM DX CPU
Introduzido	1982	1985	1988	1989
Velocidades de clock	6–12,5 MHz	16–33 MHz	16–33 MHz	25–50 MHz
Largura do barramento	16 bits	32 bits	16 bits	32 bits
Número de transistores	134.000	275.000	275.000	1,2 milhão
Dimensão mínima da tecnologia de fabricação (μm)	1,5	1	1	0,8–1
Memória endereçável	16 MB	4 GB	16 MB	4 GB
Memória virtual	1 GB	64 TB	64 TB	64 TB
Cache	–	–	–	8 kB

Evolução dos microprocessadores da Intel (continuação)

(c) Processadores da década de 1990

	486TM SX	Pentium	Pentium Pro	Pentium II
Introduzido	1991	1993	1995	1997
Velocidades de clock	16–33MHz	60–166 MHz	150–200 MHz	200–300 MHz
Largura do barramento	32 bits	32 bits	64 bits	64 bits
Número de transistores	1,185 milhão	3,1 milhões	5,5 milhões	7,5 milhões
Dimensão mínima da tecnologia de fabricação (μm)	1	0,8	0,6	0,35
Memória endereçável	4 GB	4 GB	64 GB	64 GB
Memória virtual	64 TB	64 TB	64 TB	64 TB
Cache	8kB	8kB	512 kB L1 e 1 MB L2	512 kB L2

Evolução dos microprocessadores da Intel (continuação)

(d) Processadores recentes

	Pentium III	Pentium 4	Core 2 Duo	Core 2 Quad
Introduzido	1999	2000	2006	2008
Velocidades de clock	450–660 MHz	1,3–1,8 GHz	1,06–1,2 GHz	3 GHz
Largura do barramento	64 bits	64 bits	64 bits	64 bits
Número de transistores	9,5 milhões	42 milhões	167 milhões	820 milhões
Dimensão mínima da tecnologia de fabricação (nm)	250	180	65	45
Memória endereçável	64 GB	64 GB	64 GB	64 GB
Memória virtual	64 TB	64 TB	64 TB	64 TB
Cache	512 KB L2	256 KB L2	2 MB L2	6 MB L2

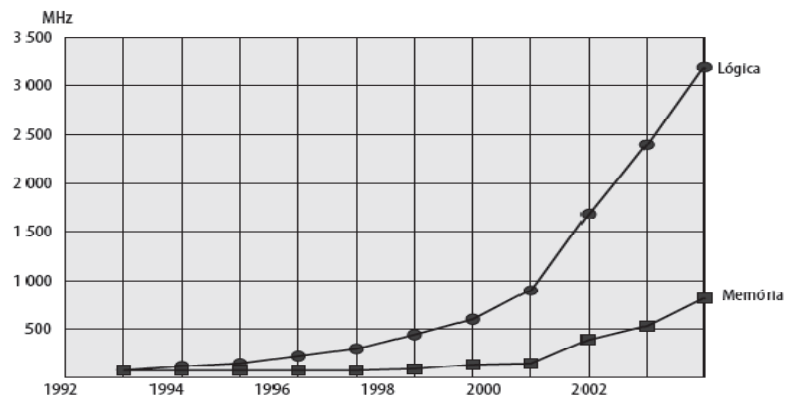
Ganhando velocidade

- Pipelining.
- Cache na placa.
- Cache L1 & L2 na placa.
- Previsão de desvio (antecipação da busca de instruções de desvio na memória)
- Análise de fluxo de dados (análise de dependência de instruções cria uma sequência otimizada)
- Execução especulativa (execução de instruções com base na previsão e análise)

Balanço do desempenho

- Aumento da velocidade do processador.
- Aumento da capacidade de memória.
- Velocidade da memória fica para trás, em relação a velocidade do processador.

Diferença de desempenho entre lógica e memória



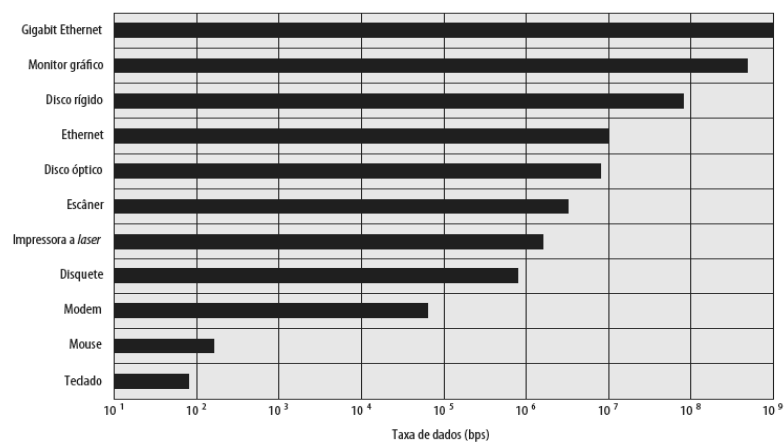
Soluções

- Aumentar número de bits recuperados de uma só vez.
 - Tornar DRAM “mais larga” ao invés de “mais profunda”.
- Mudar interface da DRAM.
 - Cache.
- Reduzir frequência de acesso à memória.
 - Cache mais complexa e cache no chip.
- Aumentar largura de banda de interconexão.
 - Barramentos de alta velocidade.
 - Hierarquia de barramentos.

Dispositivos de E/S

- Periféricos com demandas intensas de E/S.
- Grandes demandas de vazão de dados.
- Processadores podem tratar disso.
- Problema de movimentar dados.
- Soluções:
 - Caching.
 - Buffering.
 - Barramentos de interconexão de maior velocidade.
 - Estruturas de barramentos mais elaboradas.
 - Configurações de processador múltiplo.

Taxas de dados típicas dos dispositivos de E/S



A chave é o equilíbrio

- Componentes do processador.
- Memória principal.
- Dispositivos de E/S.
- Estrutura de interconexão.

Melhorias na organização e na arquitetura do chip

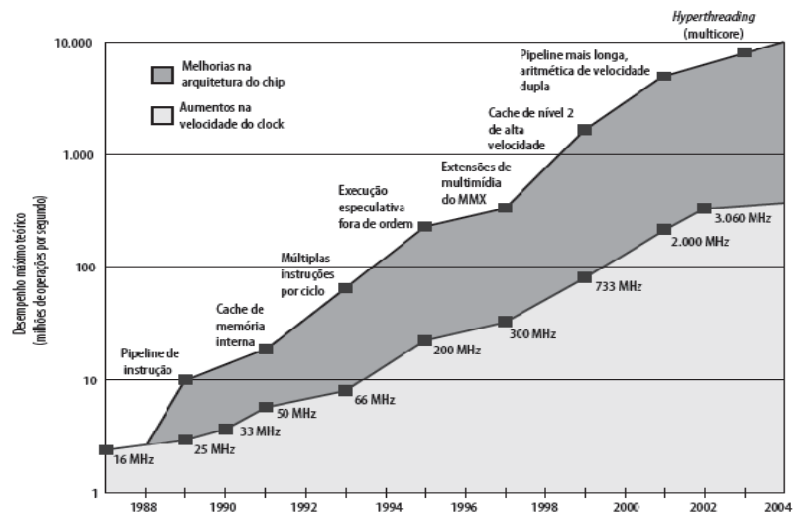
- Aumentar velocidade de hardware do processador.
 - Deve-se fundamentalmente ao encolhimento do tamanho das portas lógicas no chip.
 - Mais portas, reunidas mais de perto, aumentando a taxa de clock.
 - Redução no tempo de propagação dos sinais.
- Aumentar tamanho e velocidade das caches.
 - Dedicando à cache uma parte do chip do processador .
 - Tempos de acesso à cache caem significativamente.
- Mudar organização e arquitetura do processador.
 - Aumenta velocidade de execução efetiva.
 - Paralelismo.

Problemas com velocidade do clock e densidade da lógica

- Potência
 - Densidade de potência aumenta com densidade da lógica e velocidade do clock.
 - Dissipação de calor.
- Atraso de RC (Resistência e Capacitância)
 - Velocidade em que os elétrons fluem pela resistência e capacitância dos fios de metal que os conecta.
 - Aumentos de atraso à medida que o produto RC aumenta.
 - Interconexões de fio mais finas, aumentando a resistência.
 - Fios mais próximos, aumentando a capacitância.

- Latência da memória
 - Velocidades de memória ficam atrás das velocidades de processador.
- Solução:
 - Mais ênfase em abordagens de organização e arquitetura.

Desempenho do microprocessador Intel



Aumento da capacidade de cache

- Normalmente, dois ou três níveis de cache entre processador e memória principal.
- Densidade de chip aumentada.
 - Mais memória cache no chip.
 - Acesso mais rápido à cache.
- Chip Pentium dedicou cerca de 10% da área do chip à cache.
- Pentium 4 dedica cerca de 50%.

Lógica de execução mais complexa

- Permite execução de instruções em paralelo.
- Pipeline funciona como linha de montagem.
 - Diferentes estágios de execução de diferentes instruções ao mesmo tempo ao longo do pipeline.
- Superescalar permite múltiplos pipelines dentro de um único processador.
 - Instruções que não dependem uma da outra podem ser executadas em paralelo.

Retornos decrescentes

- Complexa organização interna dos processadores.
 - Pode obter muito paralelismo.
 - Aumentos mais significativos provavelmente serão relativamente modestos.
- Benefícios da cache estão chegando ao limite.
- Aumentar taxa de clock causa o problema da dissipação de potência.
 - Alguns limites físicos fundamentais estão sendo atingidos.

Nova técnica – múltiplos cores

- Múltiplos processadores em único chip.
 - Grande cache compartilhada.
- Dentro de um processador, aumento no desempenho proporcional à raiz quadrada do aumento na complexidade.
- Se o software puder usar múltiplos processadores, dobrar o número de processadores quase dobra o desempenho.
- Assim, melhor usar dois processadores mais simples no chip ao invés de um processador mais complexo.
- Com dois processadores, caches maiores são justificadas.
 - Consumo de potência da lógica de memória menor que lógica do processamento.

Evolução do x86

- 8080
 - Primeiro microprocessador de uso geral.
 - Caminho de dados de 8 bits.
 - Usado no primeiro computador pessoal – Altair.
- 8086 – 5MHz – 29.000 transistores.
 - Muito mais poderoso.
 - 16 bits.
 - Cache de instruções, pré-busca poucas instruções.
 - 8088 (barramento externo de 8 bits) usado no primeiro IBM PC.
- 80286
 - 16 MB de memória endereçável.
 - A partir de 1MB.

- 80386
 - 32 bits.
 - Suporte para multitarefa.
- 80486
 - Cache sofisticada e poderosa, pipeline sofisticado de instrução.
 - Coprocessador matemático embutido.

- Pentium
 - Superescalar
 - Múltiplas instruções executadas em paralelo.
- Pentium Pro
 - Organização superescalar aumentada.
 - Renomeação de registrador agressiva.
 - Previsão de desvio.
 - Análise de fluxo de dados.
 - Execução especulativa.
- Pentium II
 - Tecnologia MMX.
 - Processamento de gráficos, vídeo e áudio.
- Pentium III
 - Instruções adicionais de ponto flutuante para gráficos 3D.

- Pentium 4
 - Números romanos para números arábicos.
 - Melhorias adicionais de ponto flutuante e multimídia.
- Core
 - Primeiro x86 com dual core.
- Core 2
 - Arquitetura de 64 bits.
- Core 2 Quad – 3GHz – 820 milhões de transistores
 - Quatro processadores no chip.
- Arquitetura x86 dominante **fora dos sistemas embarcados**.
- Organização e tecnologia mudaram drasticamente.
- Arquitetura do conjunto de instruções evoluiu com compatibilidade.
- ~1 instrução acrescentada por mês durante 30 anos
- Agora existem mais de 500 instruções disponíveis.
- Veja informações detalhadas sobre processadores nas páginas Web da Intel.

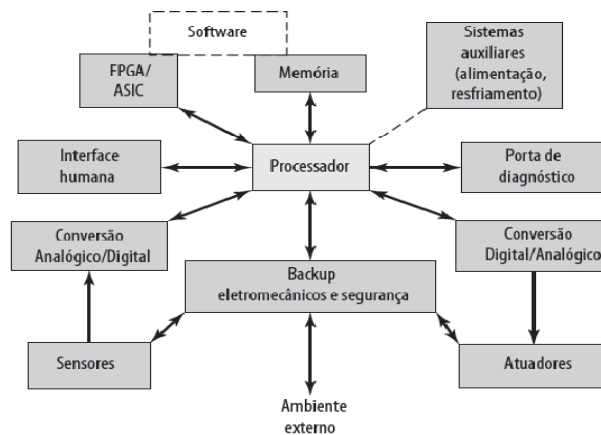
Sistemas embarcados e o ARM

- A arquitetura ARM evoluiu dos princípios de projeto RISC (capítulo 13).
- Usada principalmente em sistemas embarcados.
 - Usada dentro do produto.
 - Não para computador de uso geral.
 - Função dedicada.
 - Por exemplo, freios ABS no carro.

Requisitos dos sistemas embarcados

- Diferentes tamanhos.
 - Diferentes restrições, otimização, reuso.
- Diferentes requisitos.
 - Segurança, confiabilidade, tempo real, flexibilidade e legislação.
 - Tempo de vida
 - Condições ambientais.
 - Cargas estáticas *versus* dinâmicas.
 - Velocidades de lenta a rápida.
 - Uso intenso de computação *versus* E/S.
 - Sistemas de evento discreto até dinâmica de tempo contínuo.

Possível organização de um sistema embarcado



Evolução do ARM

- Projetado pela ARM Inc., Cambridge, Inglaterra.
- Licenciado aos fabricantes.
- Alta velocidade, pequeno tamanho do *die*, baixos requisitos de potência.
- PDAs, jogos portáteis, telefones.
—P.e., iPod, iPhone
- Acorn produziu ARM1 & ARM2 em 1985 e ARM3 em 1989.
- Acorn, VLSI e Apple Computer fundaram a ARM Ltd.

Evolução da arquitetura ARM

Tabela 2.8 Evolução da ARM

Família	Recursos notáveis	Cache	MIPS típico @ MHz
ARM1	RISC 32 bits	Nenhuma	
ARM2	Instruções de multiplicação e swap; unidade de gerenciamento de memória integrada, processador gráfico e de E/S	Nenhuma	7 MIPS @ 12 MHz
ARM3	Primeira a usar cache de processador	4 KB unificada	12 MIPS @ 25 MHz
ARM6	Primeira a aceitar endereços de 32 bits; unidade de ponto flutuante	4 KB unificada	28 MIPS @ 33 MHz
ARM7	SoC integrado	8 KB unificada	60 MIPS @ 60 MHz
ARM8	Pipeline de 5 estágios; previsão estática de desvio	8 KB unificada	84 MIPS @ 72 MHz
ARM9		16 KB/16 KB	300 MIPS @ 300 MHz
ARM9E	Instruções DSP melhoradas	16 KB/16 KB	220 MIPS @ 200 MHz
ARM10E	Pipeline de 6 estágios	32 KB/32 KB	
ARM11	Pipeline de 9 estágios	Variável	740 MIPS @ 665 MHz
Cortex	Pipeline superescalar de 13 estágios	Variável	2 000 MIPS @ 1 GHz
XScale	Processador de aplicações; pipeline de 7 estágios	32 KB/32 KB L1 512KB L2	1 000 MIPS @ 1,25 GHz

DSP = processador de sinal digital (do inglês *digital signal processor*)

SoC = sistema em um chip (do inglês *system on a chip*)

Categorias de sistemas de ARM

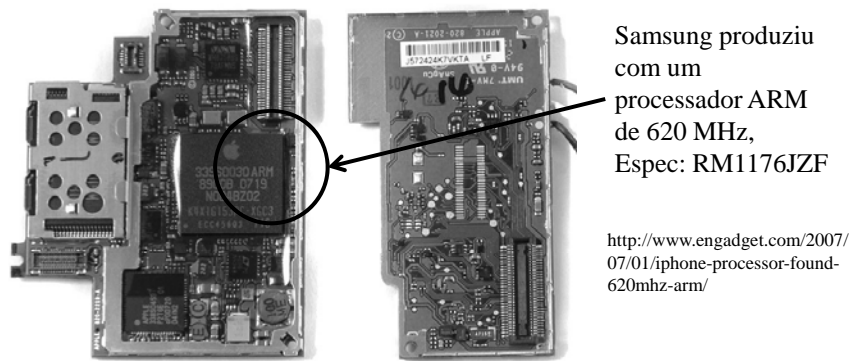
- Embarcados de tempo real
 - automotivas, de armazenamento, industriais e de redes
- Plataformas de aplicação
 - Dispositivos executando sistemas operacionais abertos, incluindo Linux, Palm OS, Symbian OS, Windows mobile.
- Aplicações seguras
 - Smart cards, terminais de pagamento

Exemplo de sistemas com ARM => iPhone



iPhone Primeira Geração

Exemplo de sistemas de ARM => iPhone



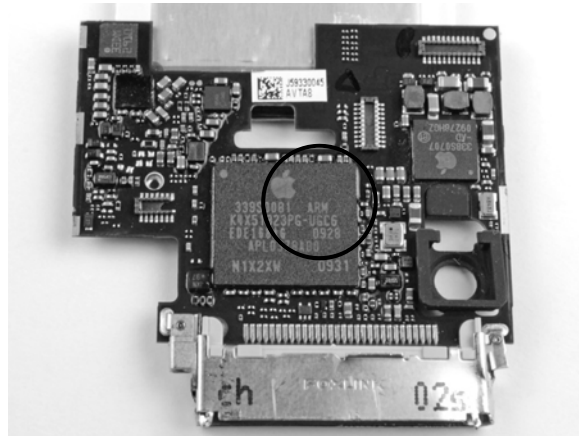
iPhone Primeira Geração

Exemplo de sistemas de ARM => iPod



iPod Nano 5ª Geração

Exemplo de sistemas de ARM => iPhone

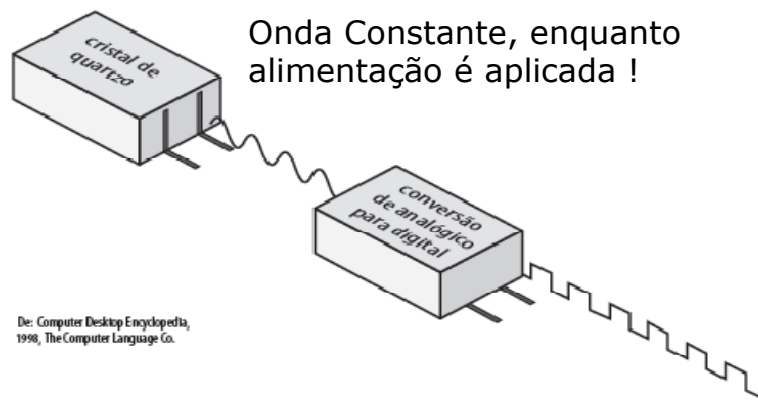


iPod Nano 5ª Geração

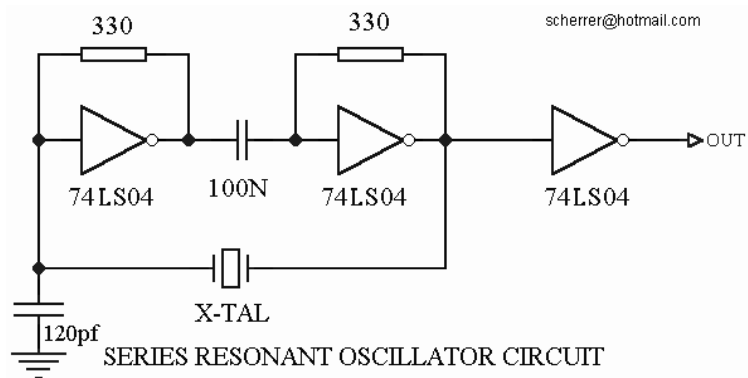
Avaliação de desempenho

- Principais parâmetros:
 - Desempenho, custo, tamanho, segurança, confiabilidade, consumo de energia.
- Velocidade do clock do sistema:
 - Em Hz ou múltiplos
 - Velocidade de clock, ciclo de clock, tempo de ciclo.
- Sinais na CPU levam tempo para se estabilizarem em 1 ou 0.
- Sinais podem mudar em diferentes velocidades.
- Operações precisam ser sincronizadas.
- Execução de instrução em etapas discretas.
 - Busca, decodificação, load e store, aritmética ou lógica.
 - Geralmente requer vários ciclos de clock por instrução.
- Pipelining gera execução simultânea de instruções.
- **Assim, velocidade de clock não é tudo.**

Clock do sistema



Exemplo de Oscilador a Cristal



Sinal de Clock

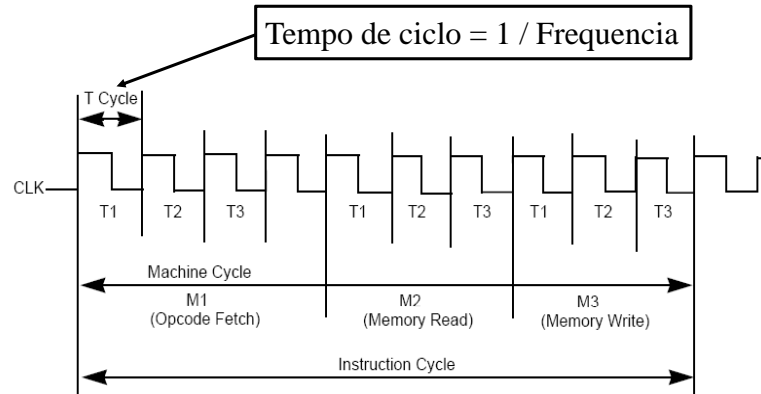


Figure 4. Basic CPU Timing Example

Taxa de execução de instrução

- f = frequência do relógio do processador
- t = ciclo constante = $1 / f$
- I_c = contagem de Instruções
 - Número de instruções executadas por um programa até que ele termine ou por um intervalo de tempo definido
- CPI = média de ciclos por instrução
- CPI_i = nr. de ciclos exigidos para a instrução tipo I
- I_i = nr de instruções executadas de tipo I
- $$CPI = \frac{\sum_{i=1,n} (CPI_i \times I_i)}{I_c}$$

Taxa de execução de instrução

- T = tempo de processador necessário para executar determinado programa pode ser expresso como:
- $T = I_c \times \text{CPI} \times t$
- Se for considerado o tempo de leitura e escrita em memória (e não somente o do processador):
- $T = I_c \times [p + (m \times k)] \times t$, onde
- p = nr de ciclos de processador
- m = nr de referências à memória
- k = razão entre o tempo de ciclo da memória e tempo de ciclo do processador

Taxa de execução de instrução

- Os 5 fatores de desempenho da equação anterior são influenciados por 4 atributos do sistema:

Tabela 2.9 Fatores de desempenho e atributos do sistema

	I_c	p	m	k	τ
Arquitetura do conjunto de instruções	X	X			
Tecnologia do compilador	X	X	X		
Implementação do processador		X			X
Hierarquia da cache e da memória				X	X

Taxa de execução de instrução

- Taxa MIPS = $I_c / T \times 10^6 = f / \text{CPI} \times 10^6$
- Supondo um processador de 400 MHz que executou 2 milhões de instruções e obteve a seguinte tabela:

Tabela 2.10 Mistura de instruções e CPI

Tipo de instrução	CPI	Número de instruções (%)
Aritmética e lógica	1	60%
Load/store com acerto de cache	2	18%
Desvio	4	12%
Referência de memória com falha de cache	8	10%

- $\text{CPI} = 0,6 + (2 \times 0,18) + (4 \times 0,12) + (8 \times 0,1) = 2,24$
- Taxa MIPS = $(400 \times 10^6) / (2,24 \times 10^6) = 178$

Taxa de execução de instrução

- Milhões de instruções por segundo (MIPS).
- Milhões de instruções de ponto flutuante por segundo (MFLOPS).
- Altamente dependente do conjunto de instruções, projeto de compilador, implementação do processador, hierarquia da memória cache e da memória principal.

Benchmarks (CISC x RISC)

$$\frac{\text{Tempo}}{\text{Programa}} = \frac{\text{tempo}}{\text{ciclo}} \times \frac{\text{ciclos}}{\text{instrução}} \times \frac{\text{instruções}}{\text{programa}}$$

- Projeto **CISC** procura diminuir as instruções/programa, em detrimento dos ciclos/instrução
- Projeto **RISC** procura diminuir o número de ciclos/instrução, em detrimento do número de instruções/programa

Benchmarks (CISC x RISC)

- Medidas como MIPS e MFLOPS provaram ser inadequadas para avaliar o desempenho dos processadores
- Exemplo: 2 máquinas que executam $A=B+C$ no mesmo tempo
 - Máquina CISC (*Complex Instruction Set Computer*)
 - $A = B + C$
 - add mem(B), mem (C), mem (A) => **1 MIPS**
 - Máquina RISC (*Complex Instruction Set Computer*)
 - $A = B + C$
 - load mem(B), reg (1)
 - Load mem(C), reg (2)
 - Add reg(1), reg (2), reg (3)
 - Store reg(3), mem (A) => **4 MIPS**

Benchmarks

- Programas elaborados para testar o desempenho.
- Escritos em linguagem de alto nível, logo portáteis.
- Representam um estilo particular de programação.
 - Sistemas, numérica, comercial.
- Facilmente medidos.
- Amplamente distribuídos.
- P.e., *System Performance Evaluation Corporation* (SPEC).
 - Pacote SPEC CPU2006 para limite de cálculo.
 - 17 programas de ponto flutuante em C, C++, Fortran.
 - 12 programas de inteiros em C, C++.
 - Mais de 3 milhões de linhas de código.
 - Métrica de velocidade e de taxa.
 - Única tarefa e vazão.

Métrica de velocidade SPEC

- Capacidade de completar uma única tarefa.
- Runtime básico para cada programa usando máquina de referência.
- Resultados são relatados como razão entre o tempo de execução de referência e o tempo de execução do sistema.
 - Tempo de execução T_{ref_i} para benchmark i na máquina de referência.
 - Tempo de execução T_{sut_i} do benchmark i na máquina de teste.

$$r_i = \frac{T_{ref_i}}{T_{sut_i}}$$

- Desempenho geral calculado pela média das razões para todos os 12 benchmarks de inteiros.
— Usa média geométrica.

$$r_G = \left(\prod_{i=1}^n r_i \right)^{1/n}$$

Métrica de taxa SPEC

- Mede vazão ou taxa de uma máquina executando uma série de tarefas.
- Múltiplas cópias de benchmarks executadas simultaneamente.
— Normalmente, mesmo número de processadores.
- Razão calculada da seguinte forma:
 - Tempo de execução de referência $Tref_i$ para benchmark i .
 - N número de cópias executadas simultaneamente.
 - $Tsut_i$ tempo decorrido desde início da execução do programa em todos os N processadores até o término de todas as cópias do programa.
 - Novamente, uma média geométrica é calculada.

$$r_i = \frac{N \times Tref_i}{T_{sut_i}}$$

Lei de Amdahl

- Gene Amdahl [AMDA67].
- *Speedup* em potencial do programa usando múltiplos processadores.
- Concluiu que:
 - Código *precisa ser paralelizável*.
 - Speedup é limitado, gerando retornos decrescentes para uso de mais processadores.
- Dependente da tarefa:
 - Servidores ganham mantendo múltiplas conexões em múltiplos processadores.
 - Bancos de dados podem ser divididos em tarefas paralelas.

Fórmula da Lei de Amdahl

- Para programa rodando em único processador:
 - Fração f do código paralelizável sem *overhead* no escalonamento.
 - Fração $(1-f)$ de código inerentemente serial.
 - T é o tempo de execução total para o programa no único processador.
 - N é o número de processadores que exploram totalmente as partes paralelas do código.

$$\begin{aligned} \text{Speedup} &= \frac{\text{tempo para executar programa em um único processador}}{\text{tempo para executar programa em } N \text{ processadores paralelos}} \\ &= \frac{T(1-f) + Tf}{T(1-f) + \frac{Tf}{N}} = \frac{1}{(1-f) + \frac{f}{N}} \end{aligned}$$

- Conclusões:
 - f pequeno, processadores paralelos têm pouco efeito.
 - $N \rightarrow \infty$, *speedup* limitado por $1/(1-f)$.
 - Retornos decrescentes para o uso de mais processadores.

Recursos da Internet

- <http://www.intel.com/>
—Procure por Intel Museum
- <http://www.ibm.com>
- <http://www.dec.com>
- Charles Babbage Institute
- PowerPC
- Intel Developer Home

Referências

- AMDA67 Amdahl, G. "Validity of the Single-Processor Approach to Achieving Large-Scale Computing Capability", *Proceedings of the AFIPS Conference*, 1967.