



Evolução e Desempenho

ENIAC – histórico



- Electronic Numerical Integrator And Computer;
- Tabelas de trajetória para armas;
- Iniciou em 1943;
- Terminou em 1946;
- Usado até 1955.

ENIAC



- Decimal;
- 20 acumuladores de 10 dígitos;
- Programado manualmente por chaves;
- 18 000 válvulas;

ENIAC



- 30 toneladas;
- 140 kW de consumo de potência;
- 5 000 adições por segundo.

von Neumann/Turing

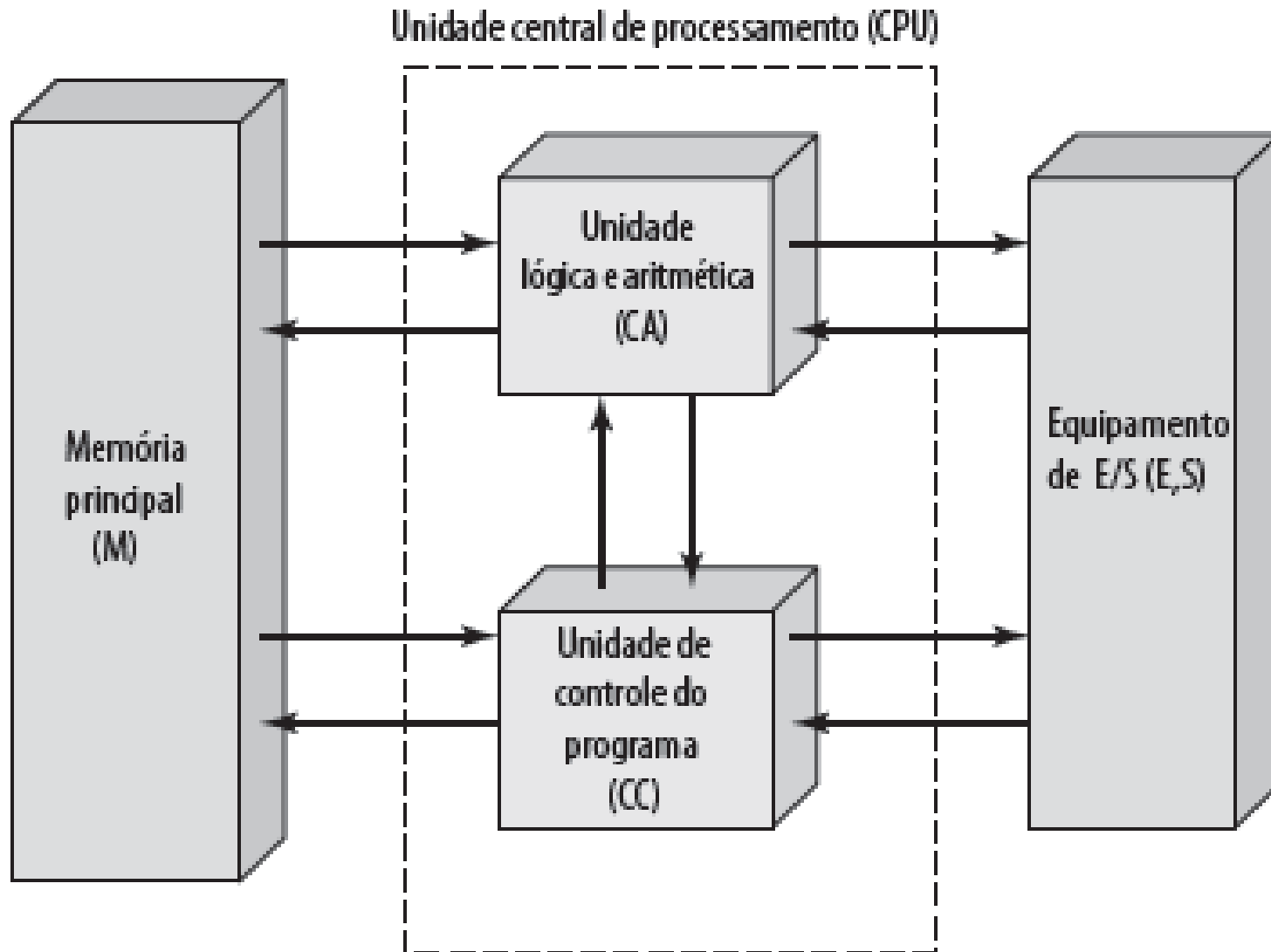


- Conceito de programa **armazenado**;
- Memória principal armazenando programas e dados;
- ULA operando sobre dados binários;



- 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.

Máquina de von Neumann



Fonte: Arquitetura e Organização de Computadores – William Stallings

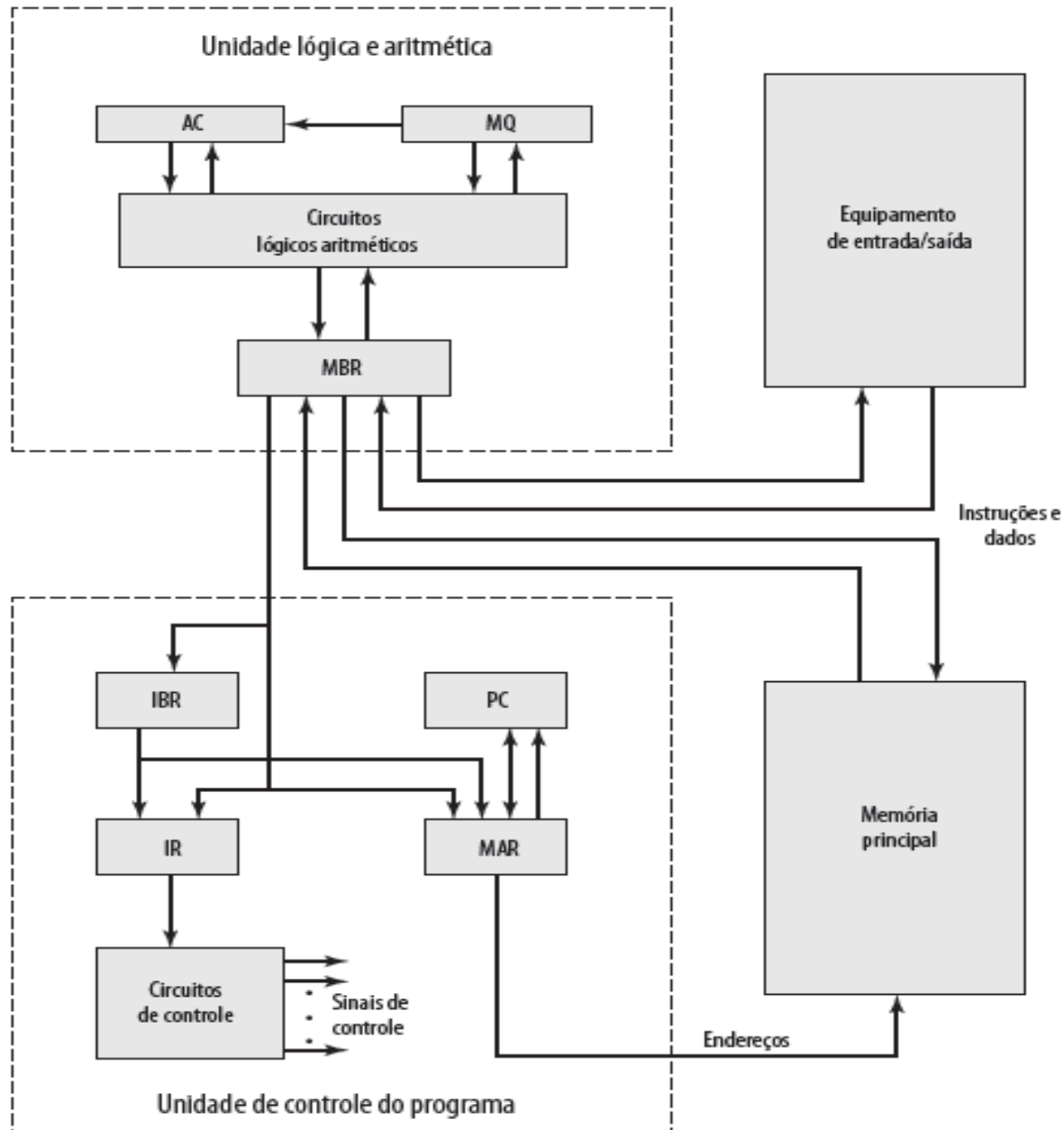


- 1000 “palavras” de 40 bits;
 - ✓ 2 instruções de 20 bits;
- Conjunto de registradores (armazenamento em UCP);
 - ✓ Registrador de buffer de memória;
 - ✓ Registrador de endereço de memória;
 - ✓ Registrador de instrução;
 - ✓ Registrador de buffer de instrução;
 - ✓ Contador de programa;
 - ✓ Acumulador;
 - ✓ Quociente multiplicador.



- 1000 “palavras” de 40 bits.
 - 2 instruções de 20 bits.
- Conjunto de registradores (armazenamento em CPU).
 - Registrador de buffer de memória.
 - Registrador de endereço de memória.
 - Registrador de instrução.
 - Registrador de buffer de instrução.
 - Contador de programa.
 - Acumulador.
 - Quociente multiplicador.

Estrutura do IAS



Fonte: Arquitetura e Organização de Computadores – William Stallings



IBM

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

Transistores



- Substituíram as válvulas;
- Menores;
- Mais baratos;
- Menos dissipação de calor;

Transistores



- Dispositivo de estado sólido;
- Feito de silício;
- Criado em 1947 (Bell Labs);



- Dispositivo composto de portas, células de memória e interconexões;
- Estas podem ser fabricadas em um semicondutor.

Gerações de computadores



- Válvula – 1946-1957.
- Transistor – 1958-1964.
- Integração em pequena escala – 1965 em diante.
 - ✓ Até 100 dispositivos em um chip.
- Integração em média escala – 1971.
 - ✓ 100-3 000 dispositivos em um chip.

Gerações de computadores



- Integração em grande escala – 1971-1977.
 - ✓ 3 000 – 100 000 dispositivos em um chip.
- Integração em escala muito grande – 1978 -1991.
 - ✓ 100 000 – 100 000 000 dispositivos em um chip.
- Integração em escala ultragrande – 1991.
 - ✓ Mais de 100 000 000 dispositivos em um chip.

Lei de Moore



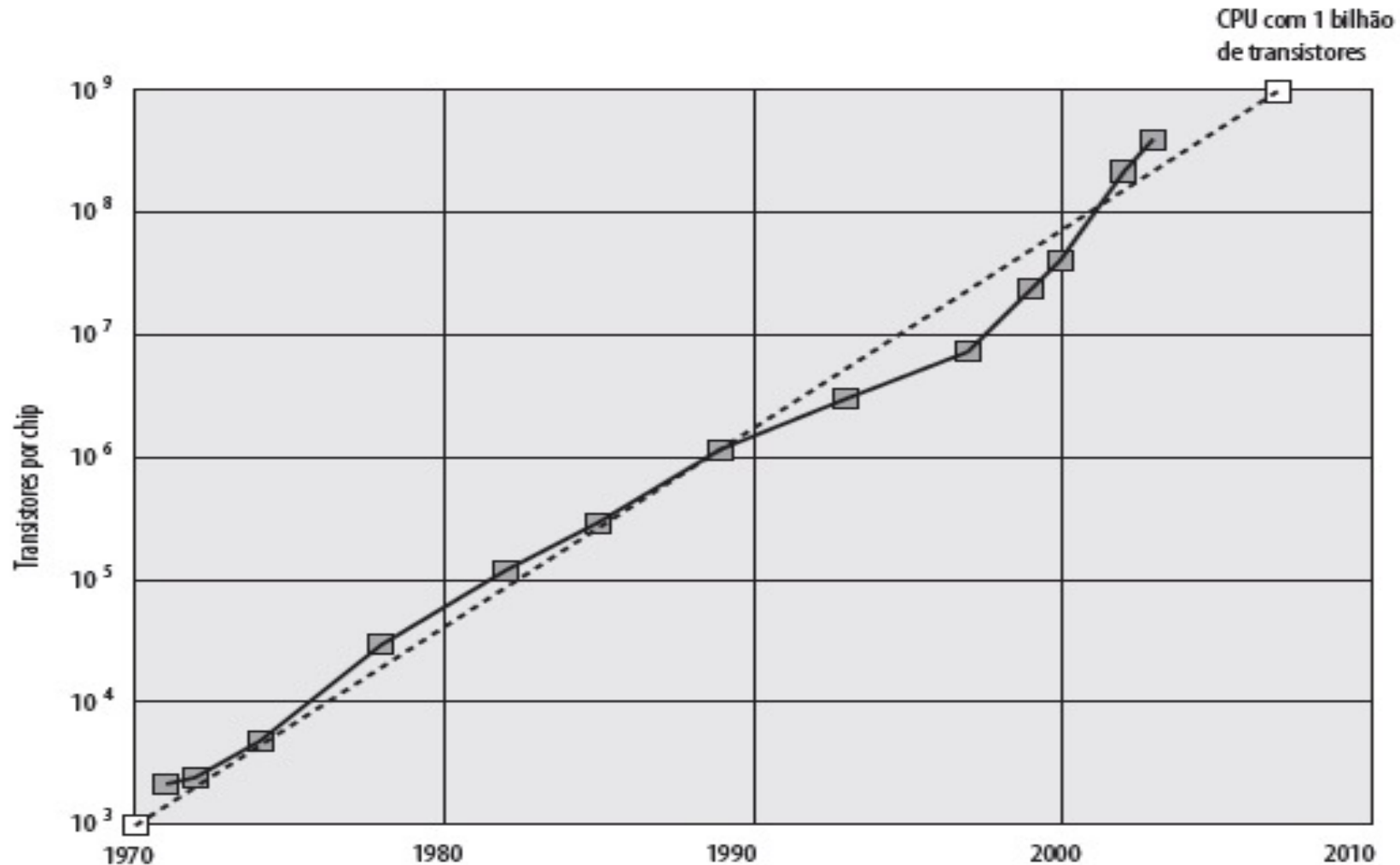
- Maior densidade de componentes no chip;
- Gordon Moore – cofundador da Intel;
- Número de transistores em um chip dobrará a cada ano;
- Desde 1970, isso diminuiu um pouco;
 - ✓ Número de transistores dobra a cada 18 meses;

Lei de Moore



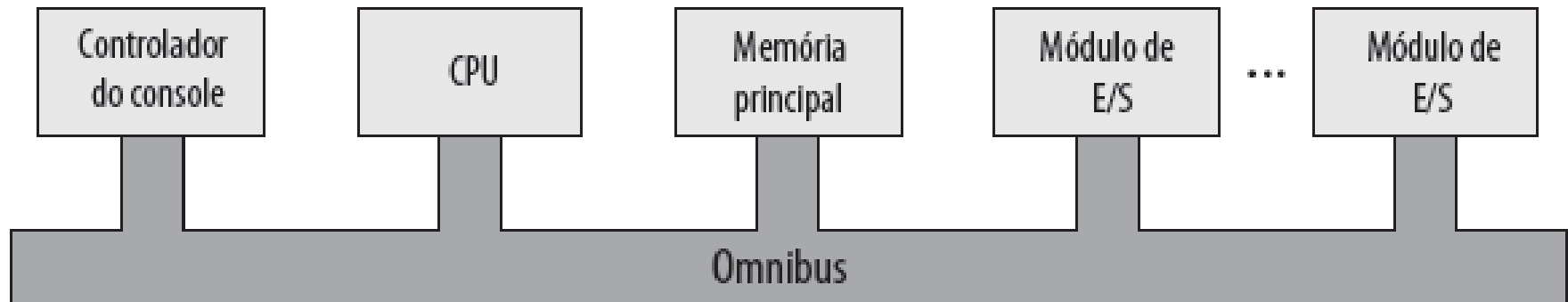
- Custo de um chip permaneceu quase inalterado.
- Maior densidade de empacotamento significa caminhos elétricos mais curtos, gerando maior desempenho.
- Menor tamanho oferece maior flexibilidade.
- Redução nos requisitos de potência e resfriamento.
- Menos interconexões aumenta a confiabilidade.

Crescimento na contagem de transistores da CPU



Fonte: Arquitetura e Organização de Computadores – William Stallings

Estrutura de barramento – Primeiro problema



Fonte: Arquitetura e Organização de Computadores – William Stallings

Memória semicondutora – Segundo Problema (1970)



- Mantém 256 bits;
- Leitura não destrutiva;
- Muito rápida;
- Capacidade dobra aproximadamente a cada ano;



- 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;

Ganhando velocidade - Intel



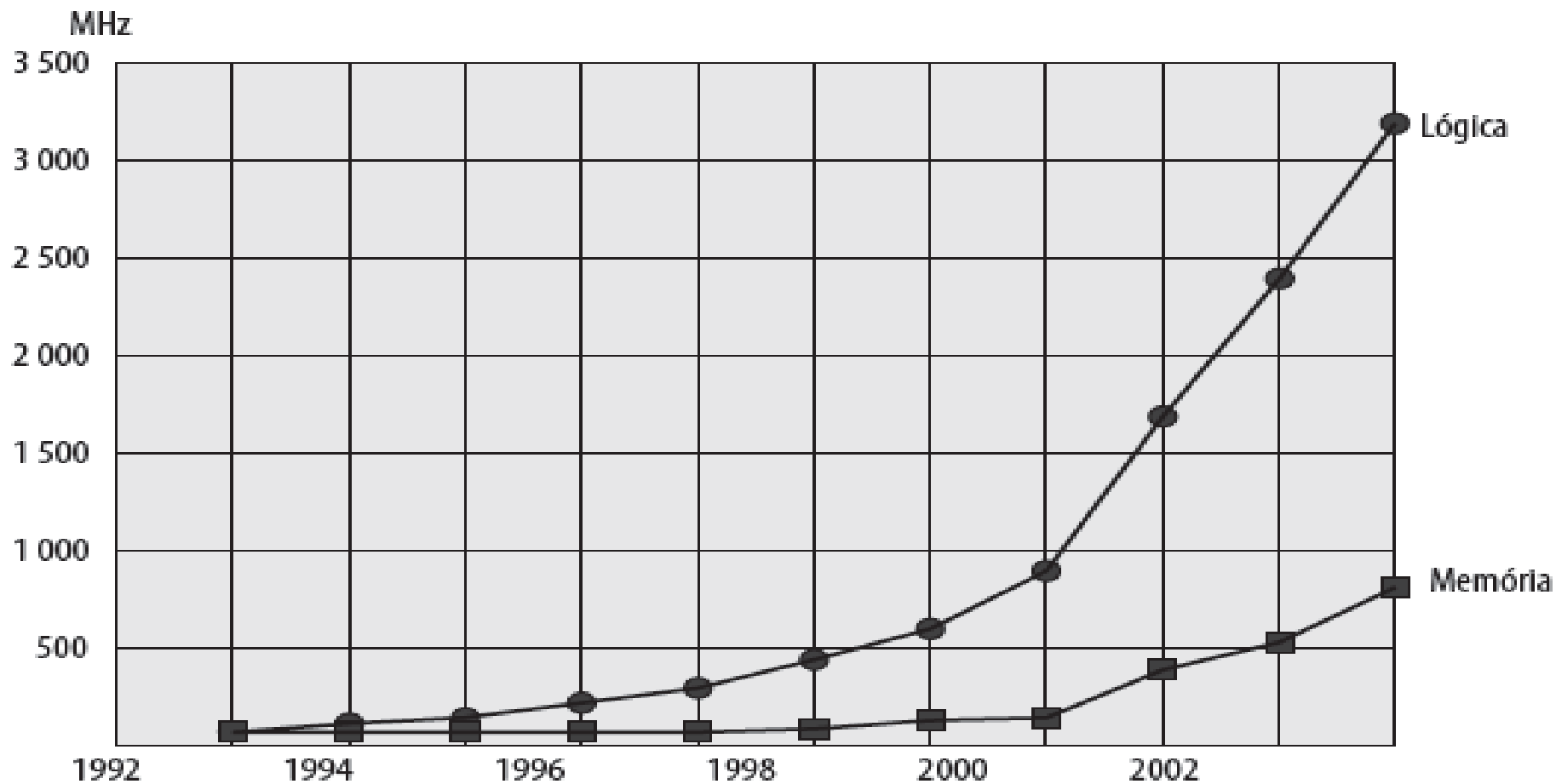
- Pipelining;
- Cache L1 & L2 na placa;
- Previsão de desvio;
- Análise de fluxo de dados;
- Execução especulativa.

Balanço do desempenho



- Aumento da velocidade do processador;
- Aumento da capacidade de memória;
- Velocidade da memória fica para trás da velocidade do processador;

Diferença de desempenho entre lógica e memória



Fonte: Arquitetura e Organização de Computadores – William Stallings

Soluções



- Aumentar número de bits recuperados de uma só vez;
- 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



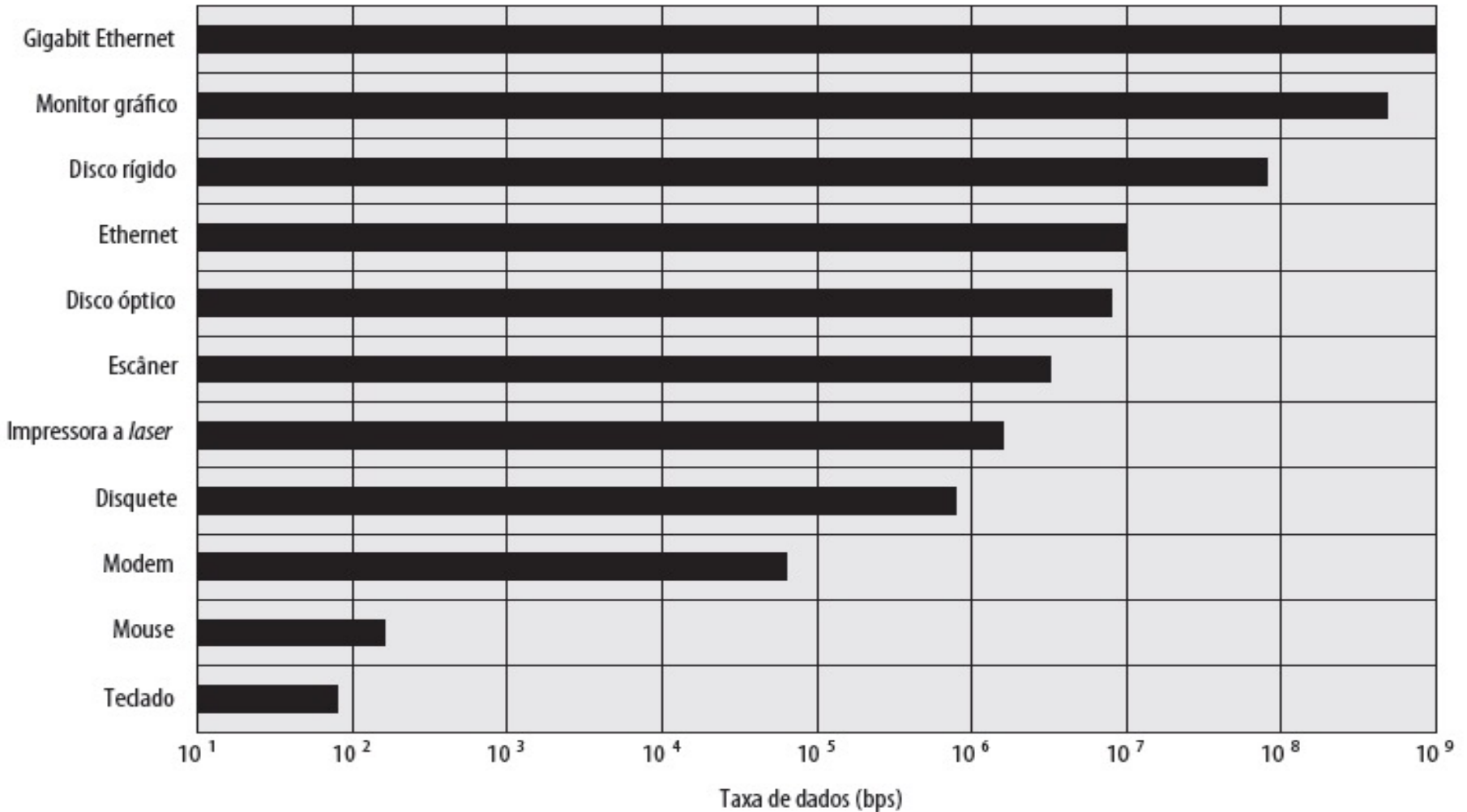
- Periféricos com demandas intensas de E/S;
- Grandes demandas de vazão de dados;
- Processadores podem tratar disso;
- Problema de movimentar dados;

Dispositivos de E/S



- 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



Fonte: Arquitetura e Organização de Computadores – William Stallings

Balanço – Importância



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

Melhorias 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.

Melhorias arquitetura do chip



- Aumentar tamanho e velocidade das caches;
 - ✓ Dedicando 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;

Problemas com velocidade do clock e densidade da lógica



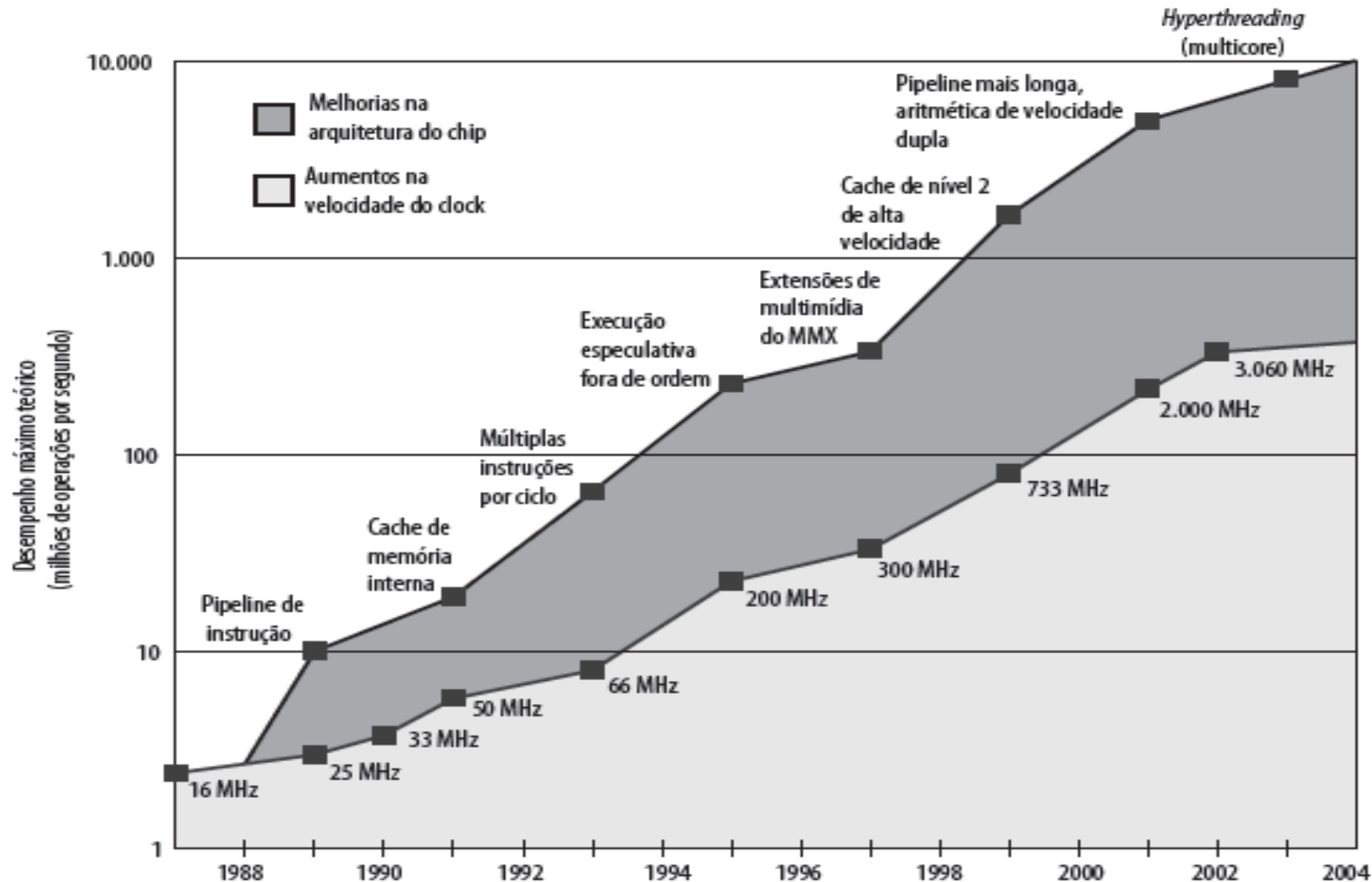
- Atraso de RC
 - ✓ Velocidade em que os elétrons fluem pela resistência e capacitância dos fios de metal que os conecta;
 - ✓ Aumentos de atraso à medida que o produto RC aumenta;
 - ✓ Interconexões de fio mais finas, aumentando a resistência;
 - ✓ Fios mais próximos, aumentando a capacitância;

Problemas com velocidade do clock e densidade da lógica



- 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



Fonte: Arquitetura e Organização de Computadores – William Stallings

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;

Aumento da capacidade de 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.

Lógica de execução mais complexa



- 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;

Nova técnica – múltiplos cores



- Assim, use 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;

Evolução do x86



- 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;

Evolução do x86



- 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.

Evolução do x86



- Pentium II
 - ✓ Tecnologia MMX;
 - ✓ Processamento de gráficos, vídeo e áudio;
- Pentium III
 - ✓ Instruções adicionais de ponto flutuante para gráficos 3D;

Evolução do x86



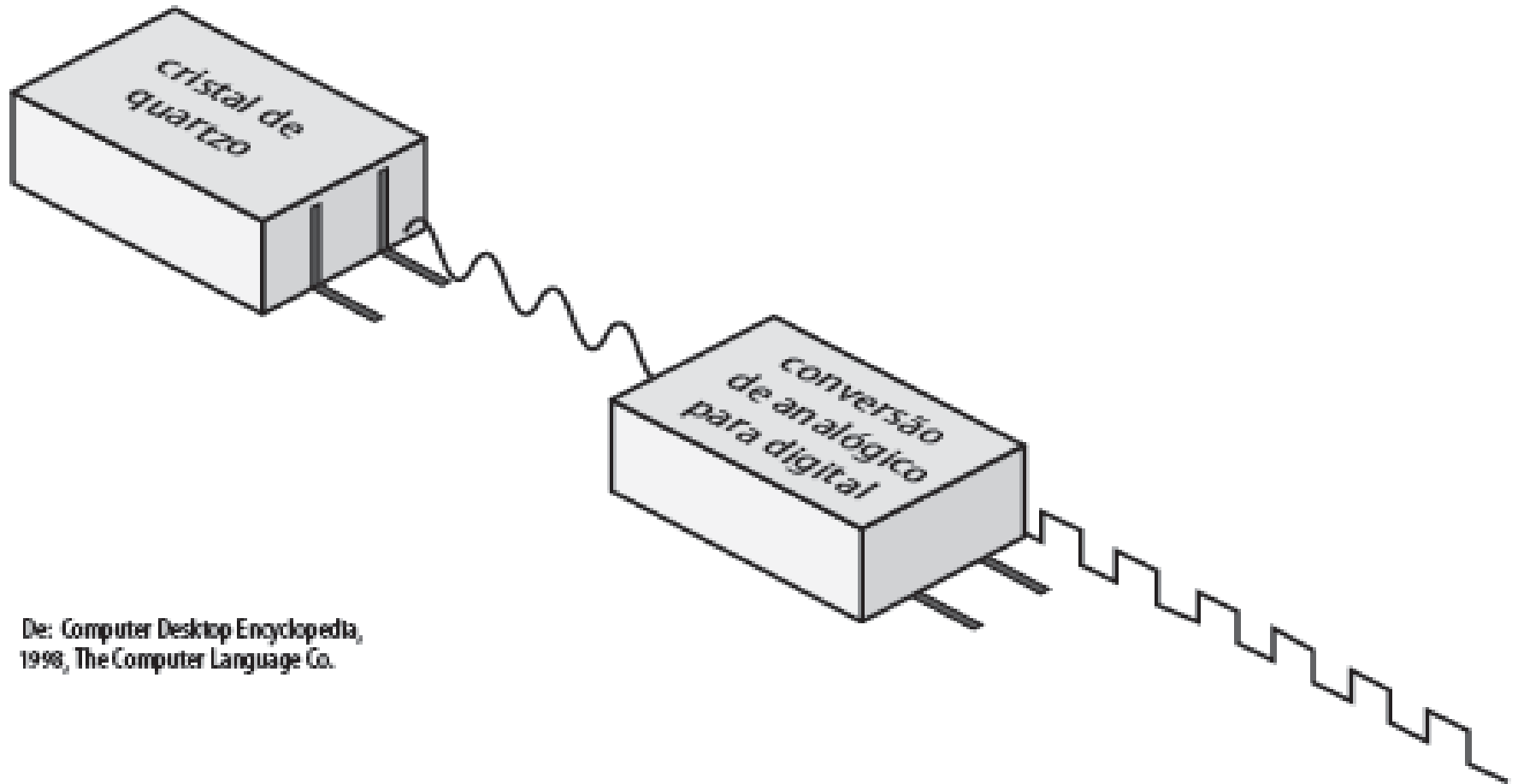
- Pentium 4
 - ✓ Números romanos para números arábicos;
 - ✓ Melhorias adicionais de ponto flutuante e multimídia;
- Core
 - ✓ Primeiro x86 com dual core;
- Core 2
 - ✓ Arquitetura de 64 bits;
- Core 2 Quad – 3GHz – 820 milhões de transistores
 - ✓ Quatro processadores no chip.

Evolução do x86



- Core i (3,5,7);
- Arquitetura x86 dominante fora dos sistemas embarcados;
- Organização e tecnologia mudaram drasticamente;
- Arquitetura do conjunto de instruções evoluiu com compatibilidade;
- 500 instruções disponíveis.

Clock do sistema



De: Computer Desktop Encyclopedia,
1998, The Computer Language Co.

Fonte: Arquitetura e Organização de Computadores – William Stallings

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 de cache e memória;

Benchmarks



- Programas elaborados para testar o desempenho;
- Escritos em linguagem de alto nível;
 - ✓ Portáteis;
- Representa o estilo da tarefa;
 - ✓ Sistemas, numérica, comercial;
- Facilmente medidos;

Lei de Amdahl



- Gene Amdahl;
- *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;

Lei de Amdahl



- 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;

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

Fórmula da Lei de Amdahl



- 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;