

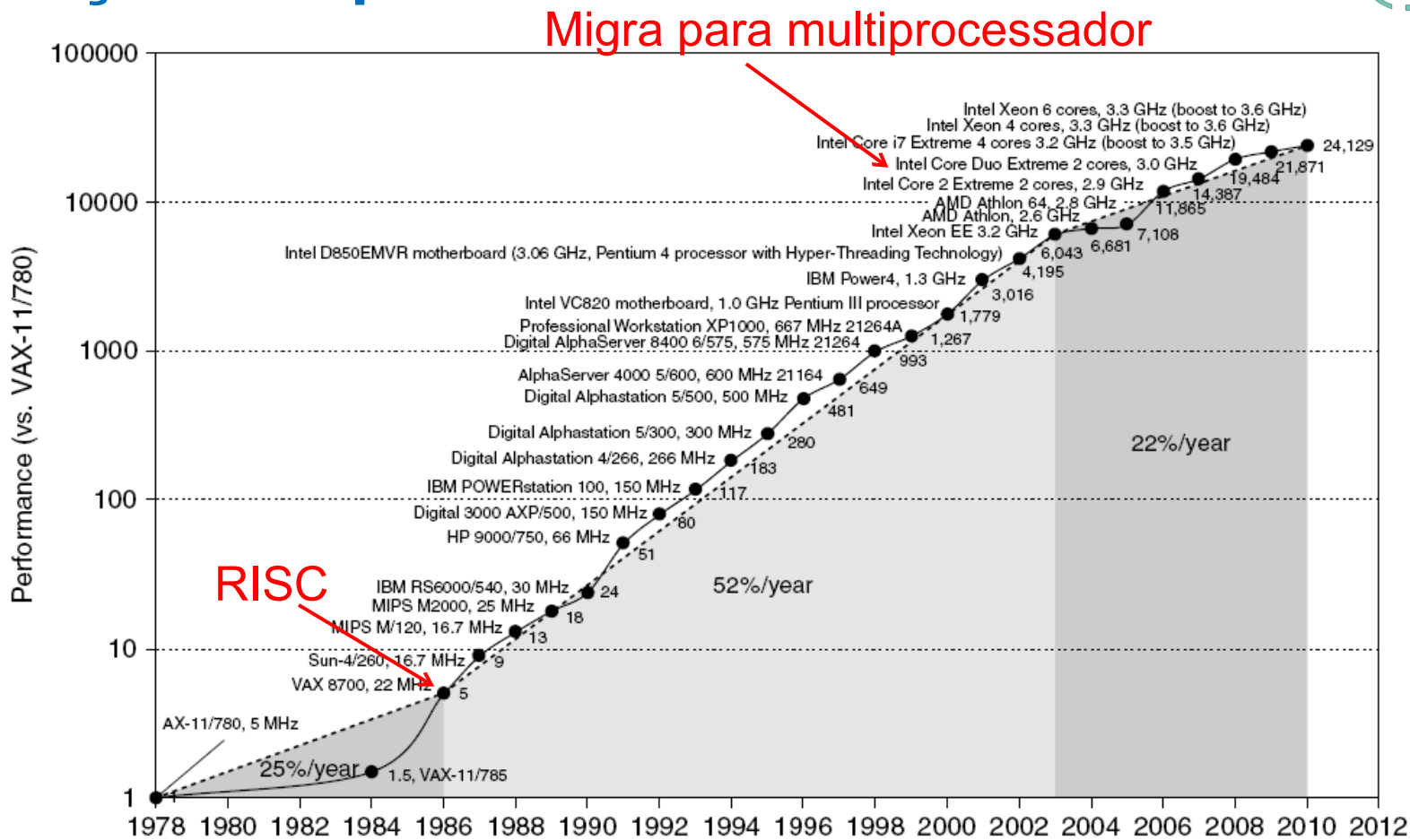
Evolução Tecnológica

GEX612 – Organização de Computadores

Prof. Luciano L. Caimi
lcaimi@uffs.edu.br

- Melhoria de performance devido:
 - Melhorias na tecnologia de semicondutores
 - Tamanho do transistor; velocidade do clock
 - Melhorias na arquitetura dos computadores
 - Microarquiteturas (ILP)
 - Compiladores e linguagens de alto nível, SOs
 - Paralelismo e interconexão

Evolução dos processadores



Transistores

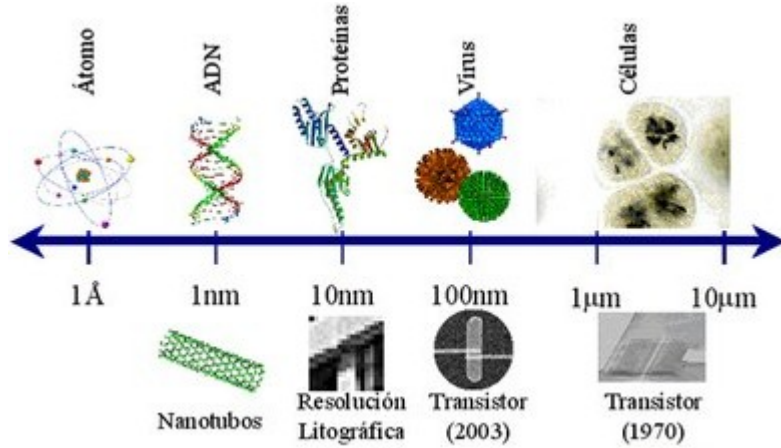
- Lei de Moore
 - 10 microns em 1971 para 0,007* microns em 2020
- Performance do transistor escala linearmente
- Densidade de integração escala quadraticamente

Fios

- Delay no fio não escala linearmente com o tamanho devido a capacitâncias parasitas

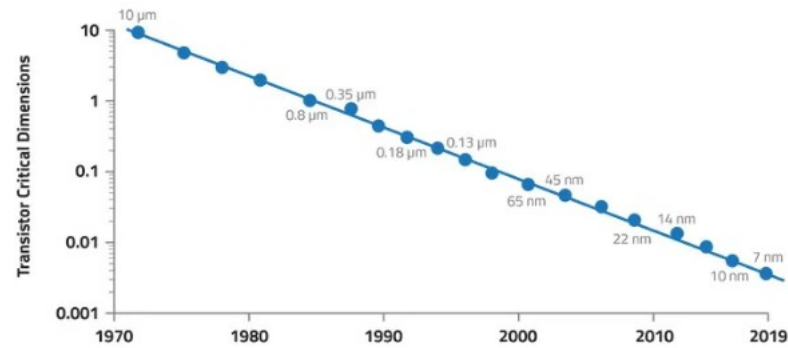
Evolução tecnológica

Transistores



$$\text{Dies per wafer} = \frac{\pi \times (\text{Wafer diameter}/2)^2}{\text{Die area}} - \frac{\pi \times \text{Wafer diameter}}{\sqrt{2} \times \text{Die area}}$$

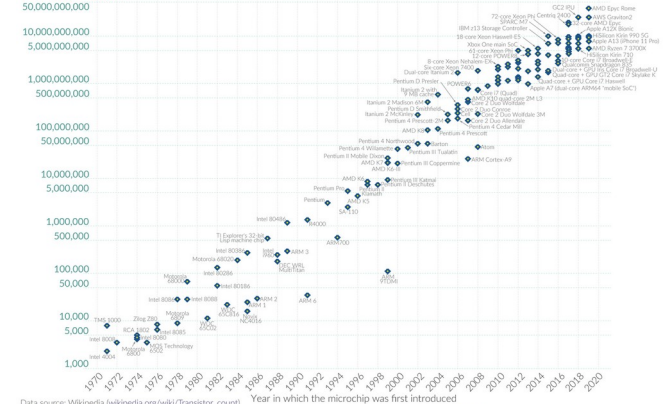
$$\text{CPU time} = \text{Instruction count} \times \text{Cycles per instruction} \times \text{Clock cycle time}$$



Moore's Law: The number of transistors on microchips doubles every two years.

Moore's law describes the empirical regularity that the number of transistors on integrated circuits doubles approximately every two years. This advancement is important for other aspects of technological progress in computing – such as processing speed or the price of computers.

Transistor count

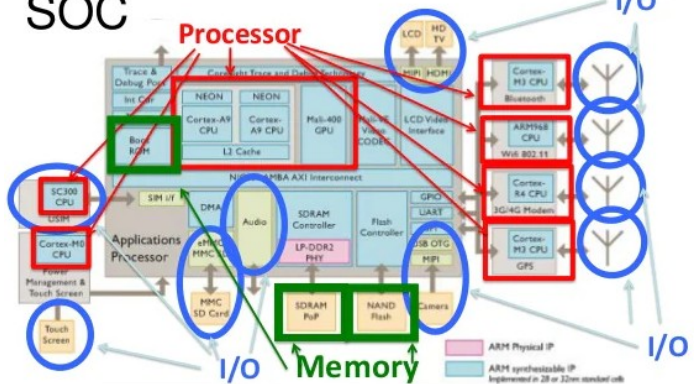


Data source: Wikipedia (wikipedia.org/wiki/Transistor_count)
OurWorldInData.org – Research and data to make progress against the world's largest problems. Licensed under CC-BY by the authors Hannah Ritchie and Max Roser.

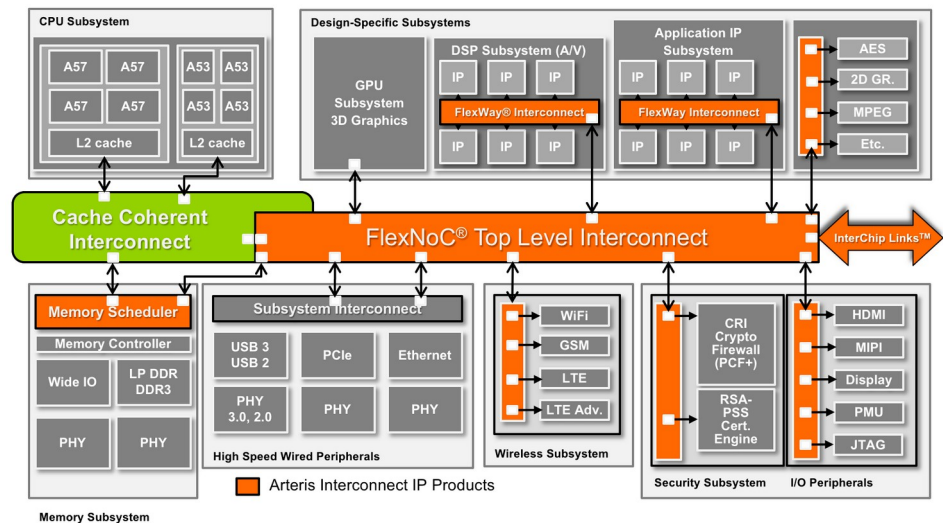
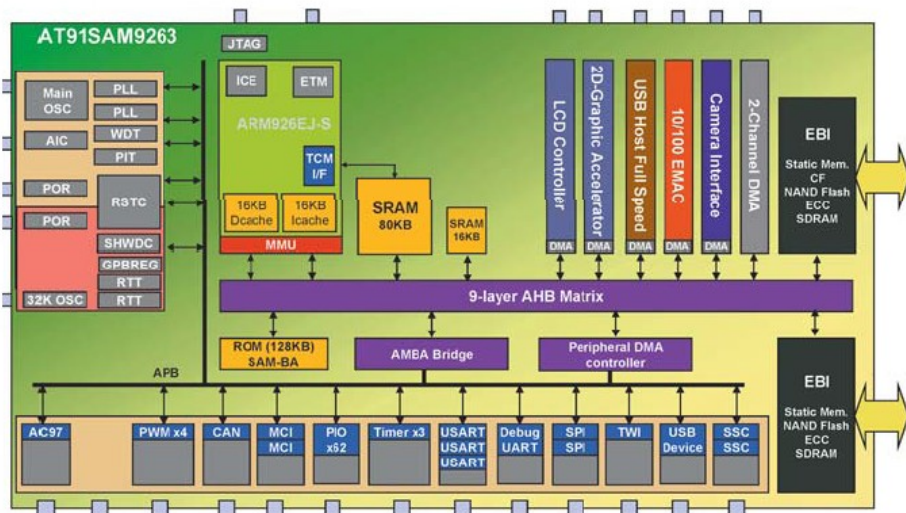
- **Dispositivos móveis pessoais (PMD)**
Consumo de energia eficiente e tempo-real (tablets, celulares, etc)
- **Desktops**
Relação preço/performance
- **Servidores**
Disponibilidade, escalabilidade e vazão (throughput)
- **Clusters / Datacentres / Comp. em nuvem**
Software como um serviço (SaaS)
Disponibilidade e relação preço/performance
Sub-classe: supercomputadores; performance de ponto flutuante
- **Computadores embarcados**

preço

Processor



Source: UC Berkeley



2020F Top 15 Semiconductor Sales Leaders (\$M, Including Foundries)

| 2020 Rank | 2019 Rank | Company | Headquarters | 2019 Total IC | 2019 Total O-S-D | 2019 Total Semi | 2020F Total IC | 2020F Total O-S-D | 2020F Total Semi | 2020/2019 % Change |
|-----------|-----------|-------------------|--------------|---------------|------------------|-----------------|----------------|-------------------|------------------|--------------------|
| 1 | 1 | Intel | U.S. | 70,797 | 0 | 70,797 | 73,894 | 0 | 73,894 | 4% |
| 2 | 2 | Samsung | South Korea | 52,486 | 3,223 | 55,709 | 56,899 | 3,583 | 60,482 | 9% |
| 3 | 3 | TSMC (1) | Taiwan | 34,668 | 0 | 34,668 | 45,420 | 0 | 45,420 | 31% |
| 4 | 4 | SK Hynix | South Korea | 22,578 | 607 | 23,185 | 25,499 | 971 | 26,470 | 14% |
| 5 | 5 | Micron | U.S. | 22,405 | 0 | 22,405 | 21,659 | 0 | 21,659 | -3% |
| 6 | 7 | Qualcomm (2) | U.S. | 14,391 | 0 | 14,391 | 19,374 | 0 | 19,374 | 35% |
| 7 | 6 | Broadcom Inc. (2) | U.S. | 15,521 | 1,722 | 17,243 | 15,362 | 1,704 | 17,066 | -1% |
| 8 | 10 | Nvidia (2) | U.S. | 10,618 | 0 | 10,618 | 15,884 | 0 | 15,884 | 50% |
| 9 | 8 | TI | U.S. | 12,812 | 839 | 13,651 | 12,275 | 813 | 13,088 | -4% |
| 10 | 9 | Infineon (3) | Europe | 7,734 | 3,404 | 11,138 | 7,438 | 3,631 | 11,069 | -1% |
| 11 | 16 | MediaTek (2) | Taiwan | 7,972 | 0 | 7,972 | 10,781 | 0 | 10,781 | 35% |
| 12 | 14 | Kioxia | Japan | 8,760 | 0 | 8,760 | 10,720 | 0 | 10,720 | 22% |
| 13 | 15 | Apple* (2) | U.S. | 8,015 | 0 | 8,015 | 10,040 | 0 | 10,040 | 25% |
| 14 | 11 | ST | Europe | 6,475 | 3,058 | 9,533 | 6,867 | 3,085 | 9,952 | 4% |
| 15 | 18 | AMD (2) | U.S. | 6,731 | 0 | 6,731 | 9,519 | 0 | 9,519 | 41% |
| — | — | Top-15 Total | | 301,963 | 12,853 | 314,816 | 341,631 | 13,787 | 355,418 | 13% |

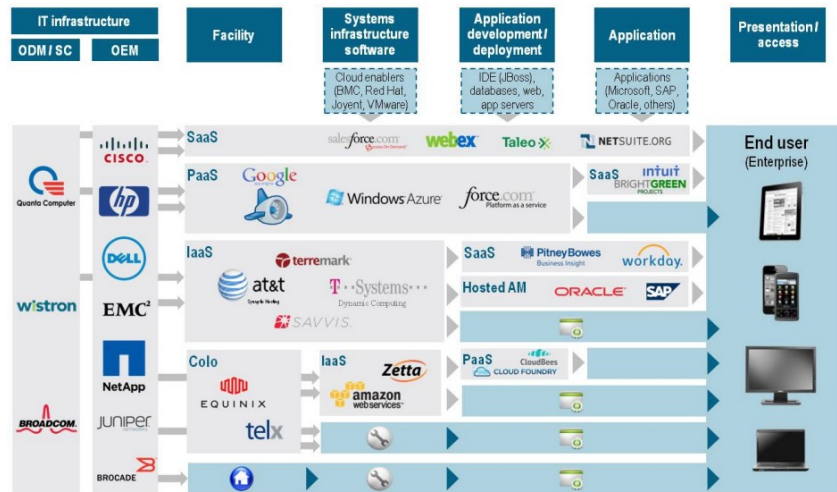
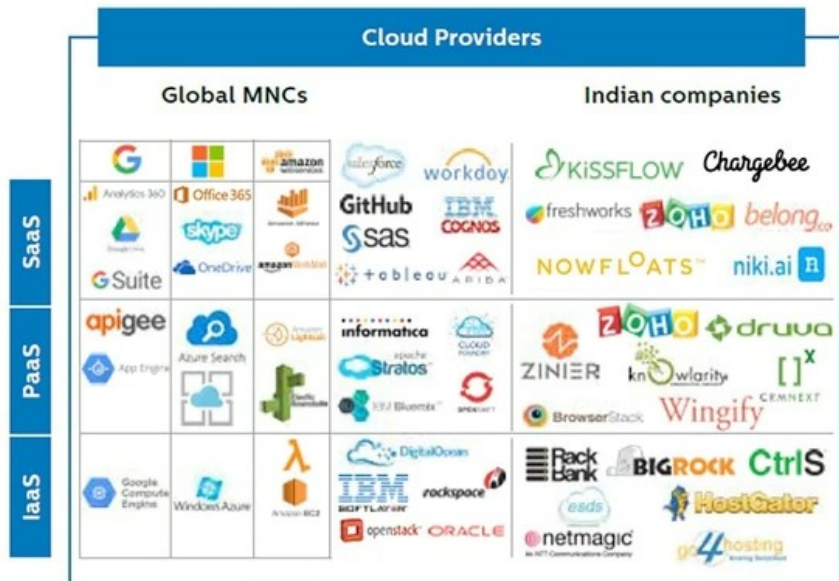
(1) Foundry (2) Fabless (3) Includes aquired company's sales in 2019 and 2020 results.

Source: Company reports, IC Insights' Strategic Reviews database *Custom processors/devices for internal

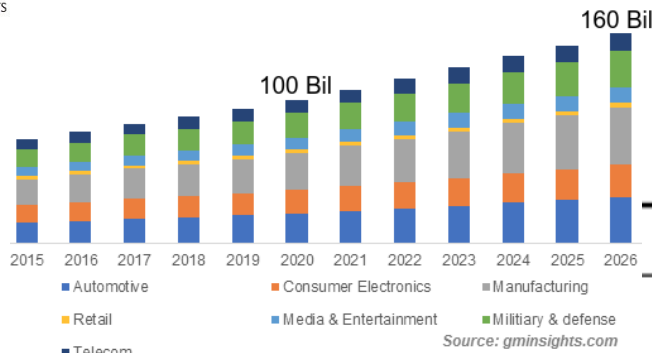
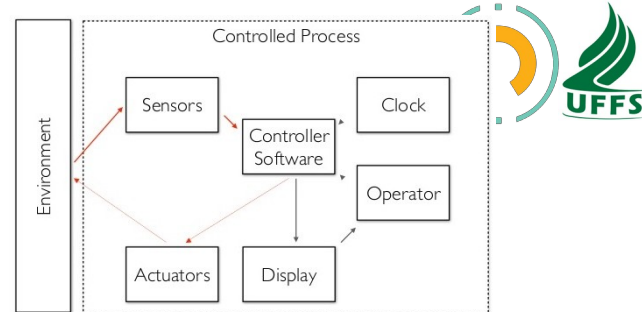
SaaS – Software como serviço

Paas – Plataforma como um serviço

IaaS – Infraestrutura como um serviço



Source: Cisco IBSG, 2012



Leading MCU Suppliers (\$M)

| 2016 Rank | Company | 2015 | 2016 | % Change | % Marketshare |
|-----------|--------------------------|-------|-------|----------|---------------|
| 1 | NXP* | 1,350 | 2,914 | 116% | 19% |
| 2 | Renesas | 2,560 | 2,458 | -4% | 16% |
| 3 | Microchip** | 1,355 | 2,027 | 50% | 14% |
| 4 | Samsung | 2,170 | 1,866 | -14% | 12% |
| 5 | ST | 1,514 | 1,573 | 4% | 10% |
| 6 | Infineon | 1,060 | 1,106 | 4% | 7% |
| 7 | Texas Instruments | 820 | 835 | 2% | 6% |
| 8 | Cypress*** | 540 | 622 | 15% | 4% |

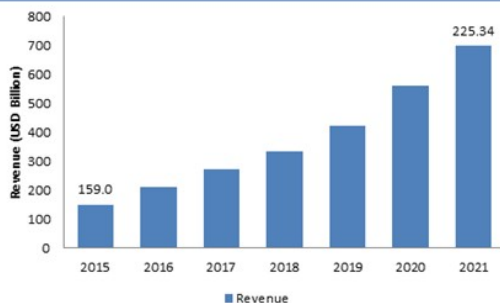
*Acquired Freescale in December 2015.

**Purchased Atmel in April 2016.

***Includes full year of sales from Spansion acquisition in March 2015.

Source: IC Insights, company reports

Global Embedded System Market Revenue, 2015- 2021 (USD Billion)



Source: Zion Research Analysis 2016

I da Fronteira Sul

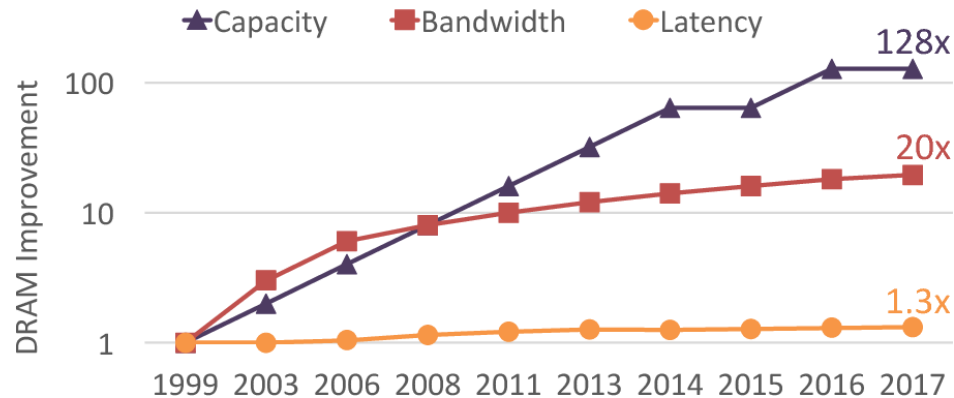
- Não há mais espaço de evolução no paralelismo a nível de instrução (Instruction Level Parallelism – ILP)
 - Aumento de performance em processadores com um único núcleo estagnou a partir de 2003
- Novos modelos para performance
 - Paralelismo a nível de dados (DLP)
 - Paralelismo a nível de threads (TLP)
 - Paralelismo a nível de requisições (RLP)

- Classes de paralelismo em aplicações:
 - Paralelismo a nível de dados (DLP)
 - Paralelismo a nível de tarefas
- Classes de paralelismo arquitetural:
 - Paralelismo a nível de Instruções (ILP)
 - Arquiteturas vetoriais / GPUs
 - Paralelismo a nível de Thread (TLP)
 - Paralelismo a nível de Requisições (RLP)

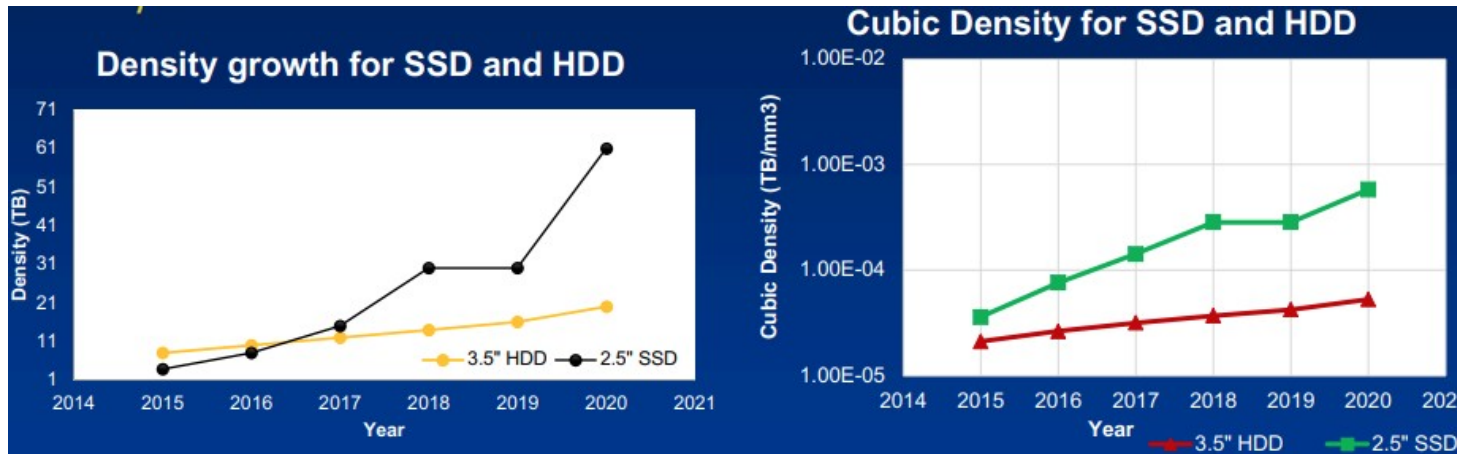
- Circuitos Integrados
 - Densidade de transistores: 35% / ano (lei de Moore?)
 - Densidade do die: 10% - 20% / ano
 - Integração geral: 40% - 55% / ano
-

Evolução tecnológica

- DRAM (memória principal)
 - Capacidade: 25% - 40% / ano
 - Latência: ??



- Flash (SSDs, pendrive, SDCards)
 - Capacidade: 50% - 60% / ano
 - Preço: 15 – 20x mais barata (por bit) que a DRAM

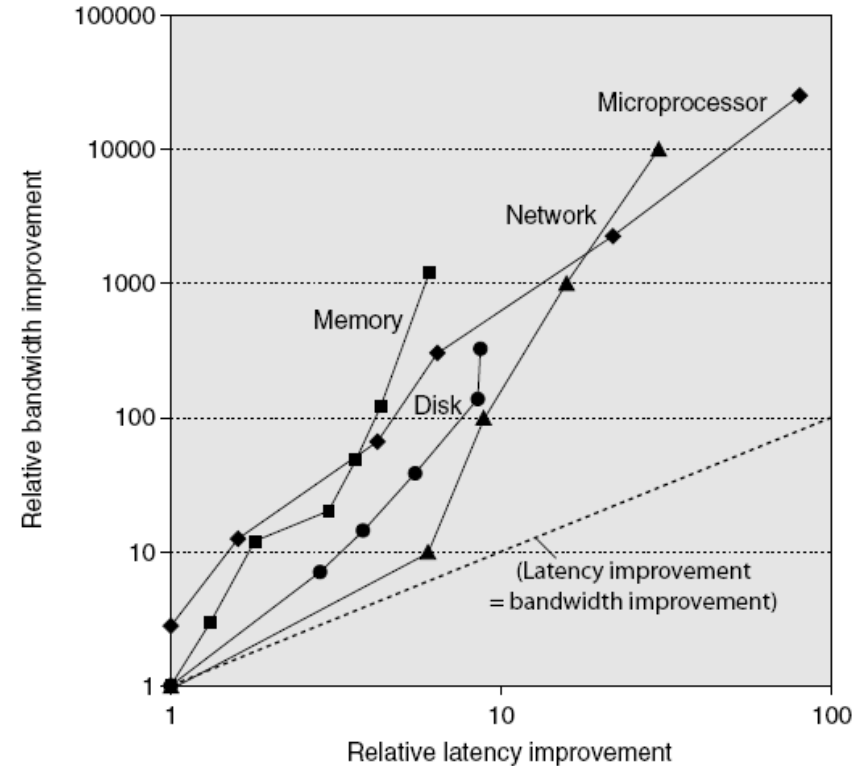


- Discos Magnéticos
 - Capacidade: 40% / ano
 - Preço: 15x – 25x mais barato (por bit) que a FLASH
 - Preço: 300x – 500x mais barato (por bit) que a DRAM

Evolução tecnológica

Largura de banda: se refere à capacidade de transmissão de dados de um ponto para outro em um determinado tempo

Latência: corresponde ao tempo decorrido entre o pedido e a entrega da informação solicitada



Evolução tecnológica: potência e energia



Problema: get power in, get power out

- potência que entra deve ser dissipada

TDP – Thermal Design Power

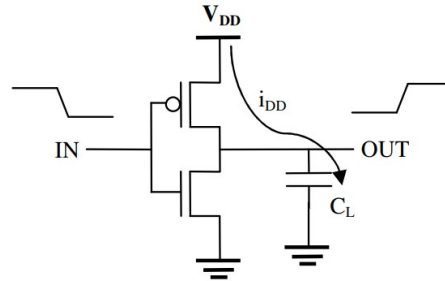
- métrica de projeto que corresponde ao consumo sustentado de energia para o qual o sistema será projetado (energia e sistema de resfriamento);
- valor menor que o pico de consumo e maior que o consumo médio;
- diminuição da frequência de clock dinamicamente para limitar o consumo

Evolução tecnológica: potência e energia



Energia dinâmica:

- Ocorre devido ao chaveamento do transistor $0 \rightarrow 1$ ou $1 \rightarrow 0$



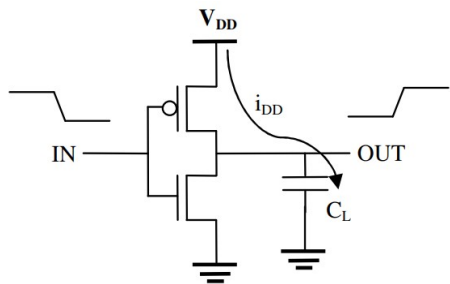
- proporcional a capacitância da carga: C_L
- quadrática com a tensão de alimentação: V_{DD}^2

Evolução tecnológica: potência e energia



Potência dinâmica:

- Ocorre devido ao chaveamento do transistor $0 \rightarrow 1$ ou $1 \rightarrow 0$



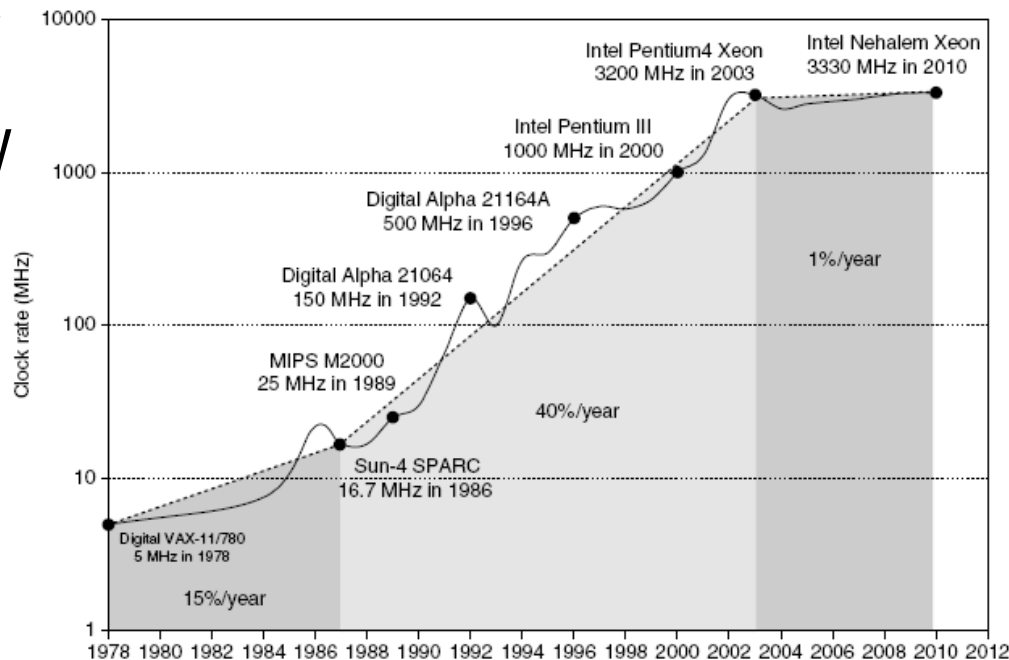
$$P_{chaveamento} = \alpha \cdot C_L \cdot V_{DD}^2 \cdot f_{clock}$$

- proporcional a frequência de chaveamento: f
- proporcional a capacitância da carga: C_L
- quadrática com a tensão de alimentação: V_{DD}^2

Reduzir a frequência reduz a potência, mas não reduz a energia

Evolução tecnológica: potência e energia

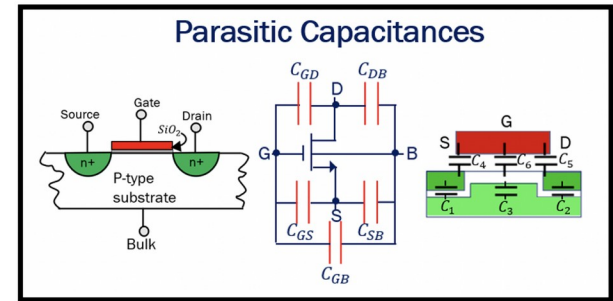
- Intel 80386 consumia ~ 2 W
- Intel Core i7 consome 130 W
- Calor tem que ser dissipado de um chip 1,5 x 1,5 cm
- Este é o limite que pode ser resfriado com ar



- Técnicas de redução de potência dinâmica
 - Desabilitar relógio de módulos inativos (clock gating)
 - Escalonamento dinâmico de voltagem/frequência (DVFS)
 - Colocar DRAMs e discos em modo baixa potência
 - Cores mistos (ARM Big.little)
 - Overclocking

Potência estática

- Corrente_{estática} x Tensão
- Escala com o número de transistor
- Devido a:
 - capacitâncias parasitas
 - distância dos fios
- Técnicas para minimizar
 - Power gating
 - Desligar módulos/cores (dark silicon)



Visão de projeto de uma arquitetura



- Visão antiga
 - Projeto do conjunto de instruções (ISA - Instruction Set Architecture)
 - Decisões visando definir: registradores; modos de endereçamento; operandos das instruções; codificação das instruções; instruções controle de fluxo, etc
- Visão atual
 - Atender requisitos específicos da máquina alvo
 - Projeto para maximizar performance considerando restrições de custo, consumo energético e disponibilidade
 - Inclui o ISA, microarquitetura e o hardware