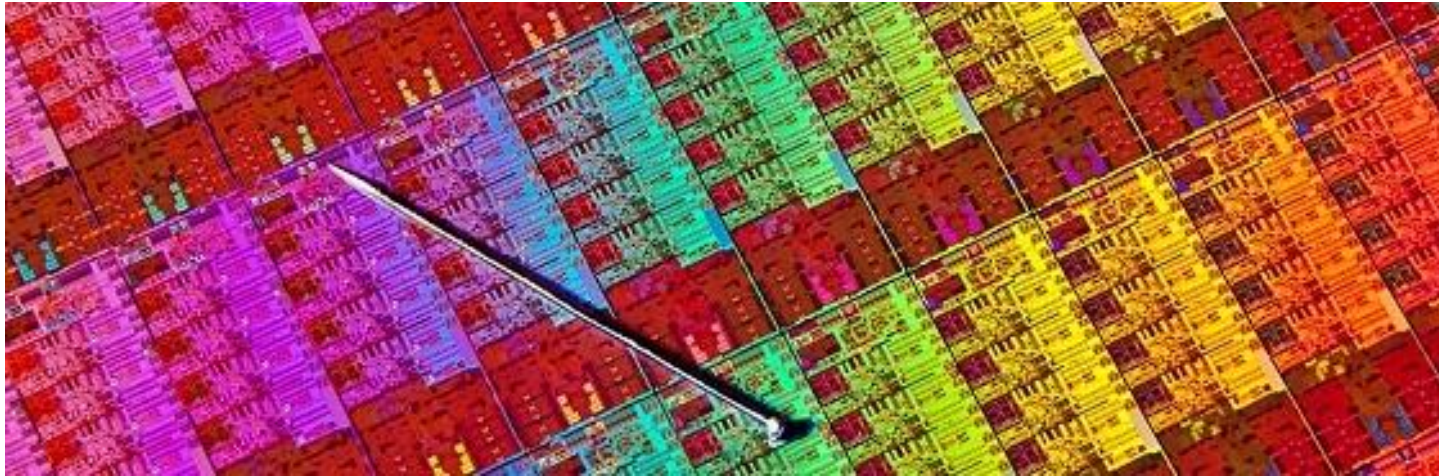




Haswell wafer

RCM00014

ARQUITETURA DE COMPUTADORES

- Prof. Luciano Bertini <lbertini@id.uff.br>
- Site:
 - <http://www.professores.uff.br/lbertini/>

Objetivos do Curso

- Entendimento mais aprofundado do funcionamento de um computador, no nível da CPU.
- O aluno aprenderá como comparar duas arquiteturas distintas sob o aspecto do desempenho e como são aplicados os benchmarks.
- Entender o funcionamento do conjunto de instruções de uma máquina e qual a influencia dele na eficiência dos compiladores.
- Paralelismo a nível de instruções: pipeline.
- O curso também aborda multiprocessadores e processamento vetorial, como visão geral de arquiteturas avançadas.

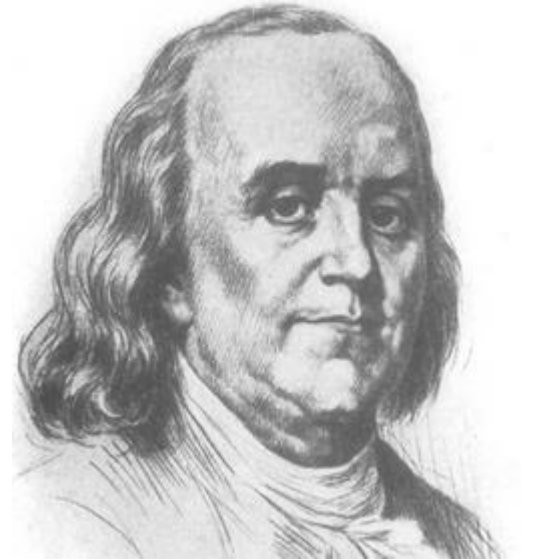
Presença:

- Mínimo de 75% de presença. (Regulamento, Art. 80, §14)
- Reprovação com média 0 se presença < 75%. (Regulamento, Art. 80, §14)
- Nenhuma falta será abonada. (Regulamento Art. 80, §15)

Avaliação:

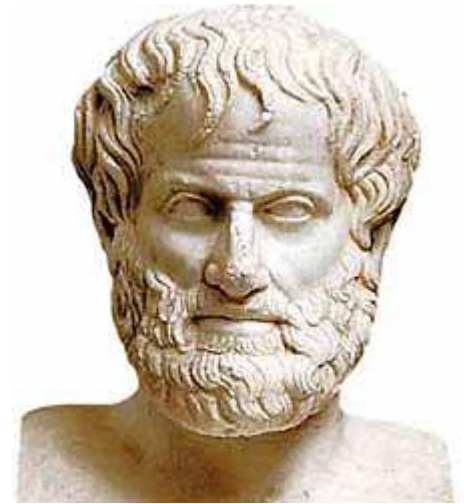
- Serão dados dois exercícios (EX1 e EX2), um trabalho prático (TP) e duas provas (P1 e P2).
- A nota final será dada por:

$$N = \frac{3P1+3P2+2TP+EX1+EX2}{10}$$



**“Diga-me e eu esquecerei, ensina-me e eu
poderei lembrar, envolva-me e eu aprenderei”**

– Benjamin Franklin



“Existe só uma maneira de se evitar as críticas: não fazer nada, não dizer nada e não ser nada”

– Aristoteles

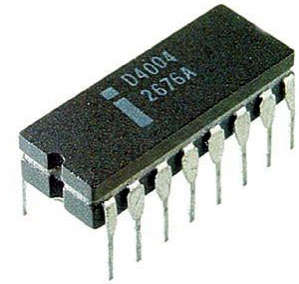
CAPÍTULO 1

INTRODUÇÃO

- Progresso da tecnologia
- Desempenho dos processadores
- Classes de Computadores
- Arquitetura do Conjunto de Instruções (ISA)
- Exemplo: MIPS
- Energia e Potência nos microprocessadores
- Potência estática vs potência dinâmica
- Processo de fabricação de chips

Progresso da tecnologia

- A tecnologia de computadores evoluiu muito desde a criação dos primeiros microprocessadores disponíveis comercialmente (Ex.: Intel 4004)
- Um computador que hoje custa menos de R\$1.000,00 tem mais desempenho, mais memória e muito mais disco que um computador que em 1985 custava R\$2.000.000,00.
- Duas forças impulsionam esse progresso: o avanço tecnológico e o avanço das *arquiteturas de computadores*



Intel 4004



Intel C4004 microprocessor

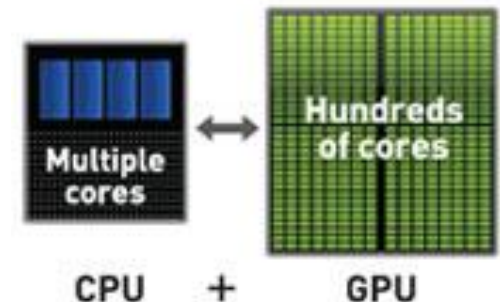
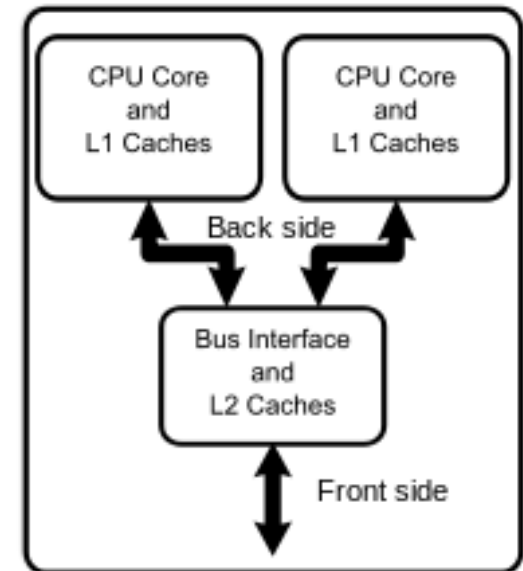
Produced	From late 1971 to 1981
Common manufacturer(s)	Intel
Max. CPU clock rate	740 kHz
Min. feature size	10µm
Instruction set	4-bit BCD-oriented
Application	Busicom calculator, arithmetic manipulation
Successor	Intel 4040 Intel 8008
Package(s)	16-pin DIP

Progresso da tecnologia

- Combinado com o baixo custo da produção em larga escala, todo o mercado de computadores passou a depender dos microprocessadores
- Duas outras características favoreceram o sucesso comercial de novas arquiteturas: eliminação do uso da linguagem *assembly* e o aparecimento de sistemas operacionais independentes e abertos, como o Linux
- Um avanço significativo em 1980 foi o aparecimento das arquiteturas RISC (*Reduced Instruction set Computer*) usando técnicas de paralelismo em nível de Instruções (ILP) com pipeline e múltiplas instruções executadas simultaneamente.

Processadores multi-core

- Componente processador único com dois ou mais UCPs.
- Processadores Multicore começaram a surgir em 2004.
- O desempenho depende muito do software, da facilidade de paralelização da aplicação. O ganho é limitado pela fração do programa que não pode ser paralelizado, como prevê a **Lei de Amdhal** (que veremos depois)
- O **speed-up** (que também veremos depois) pode chegar próximo ao número de núcleos
- Aplicações gráficas são as mais fáceis de se paralelizar: paralelismo de dados
- Processamento paralelo com GPUs
- CUDA (*Compute Unified Device Architecture*) é uma plataforma de programação paralela e modelo de programação desenvolvido pela NVIDIA [VER VÍDEO LEONARDO 2.0](#)



Processadores de smartphones



- A compra de um telefone hoje em dia exige a comparação dos processadores adotados por cada modelo
- Grandes empresas como a Samsung fabricam seus próprios modelos SoCs

Celular Vivo?

www.vivo.co.in/VSeries/

Apps SIGAC - Sistema de G Speechnotes | Ditado ★ Bookmarks Source code beautif Lição 4 - FAMÍLIAS DE Agendamento Fácil Apartamento à venda Portal de Serviços do LECTURE_4.pdf

vivo HOME PRODUCTS ABOUT FIND US NEWS ROOM SUPPORT

V5Plus
[20MP] DUAL FRONT CAMERA
Perfect Selfie

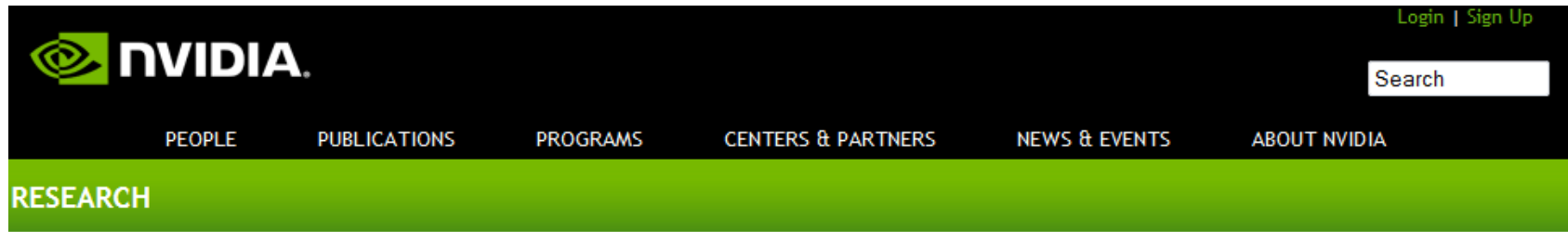
20MP+8MP Dual Front Camera Front Moonlight Glow 4GB RAM 64GB ROM

Compare Now

V Series
Best Sellers



Instituto de Computação da UFF é centro de pesquisa CUDA da NVIDIA



[NVIDIA](#) > [Research](#)

CENTERS & PARTNERS

- [CUDA Centers of Excellence](#)
- [CUDA Fellows](#)
- [CUDA Research Centers](#)
- [CUDA Teaching Centers](#)

PROGRAMS

- [CUDA Center of Excellence](#)
- [CUDA Research Center](#)
- [CUDA Teaching Center](#)
- [Graduate Fellowship](#)
- [Hardware Donation Program](#)

ADDITIONAL INFORMATION

- [CUDA Courses Around the World](#)
- [CUDA Course Materials](#)
- [CUDA Developer Curriculum](#)
- [CUDA Forums](#)
- [CUDA Zone](#)
- [Developer Zone](#)
- [GPU Technology Conference](#)
- [Intern & Coop Programs](#)

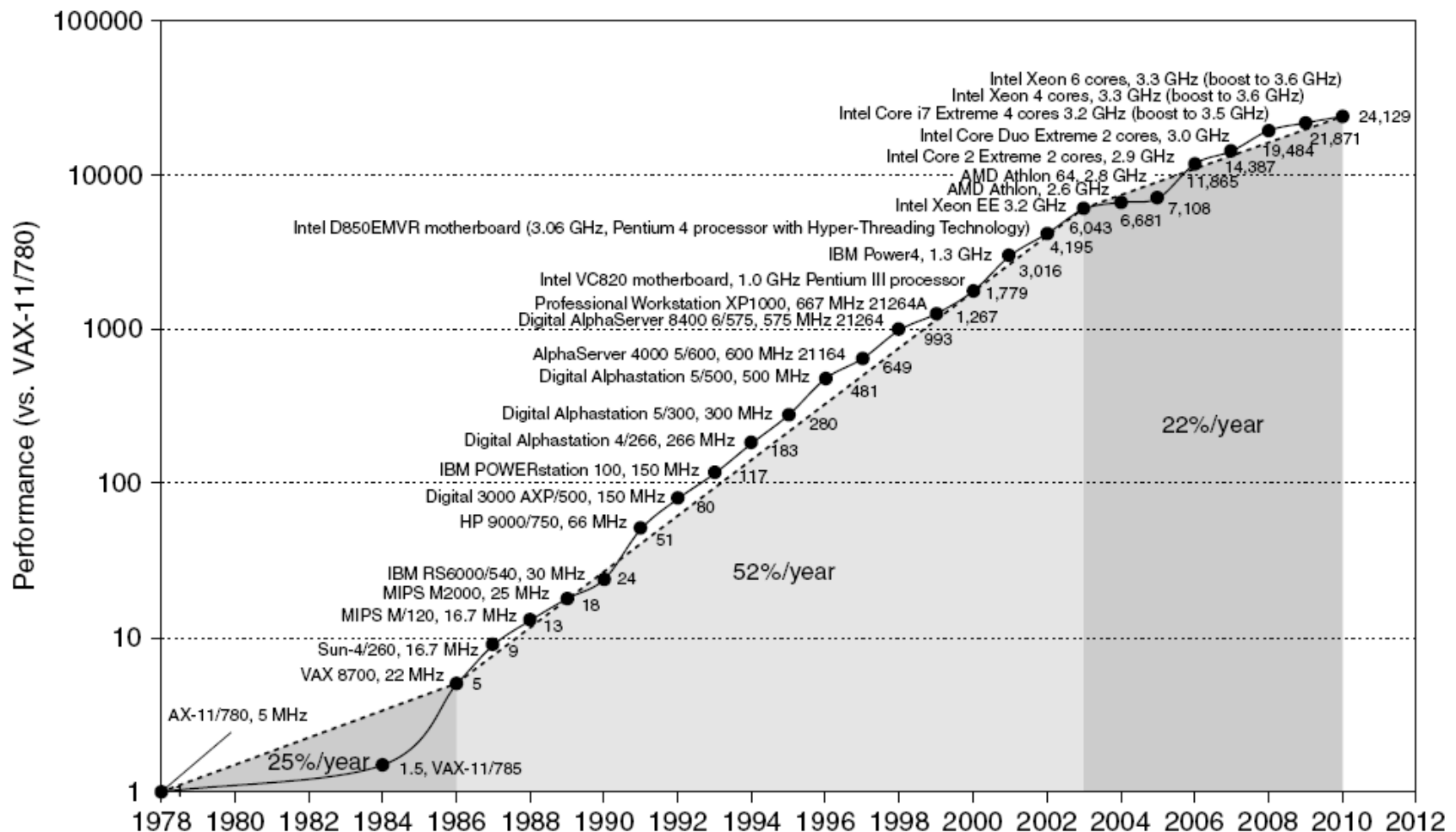
Universidade Federal Fluminense CUDA Research Center Summary



About the CUDA Research Center at Universidade Federal Fluminense

The Computer Science Department of [Universidade Federal Fluminense \(UFF\)](#) is well-recognized and offers a 5 star undergraduate and graduate Computer Science Program. These programs offer a master and doctor program in various areas of computer science. The CS department was created in 1974 and today, the department has 61 professors, many of them very well recognized in their research field and is listed among the top ten CS research centers in Brazil. Medialab@UFF is the laboratory of the Department dedicated to videogames, digital entertainment and virtual reality. The laboratory is a reference in Brazil for GPU Computing and CUDA, being recognized not only by the academy, but also by many companies. Among different projects, Medialab develops GPU Computing solutions for Oil and Gas, Naval and Entertainment industry.

Desempenho dos sistemas desde 1970



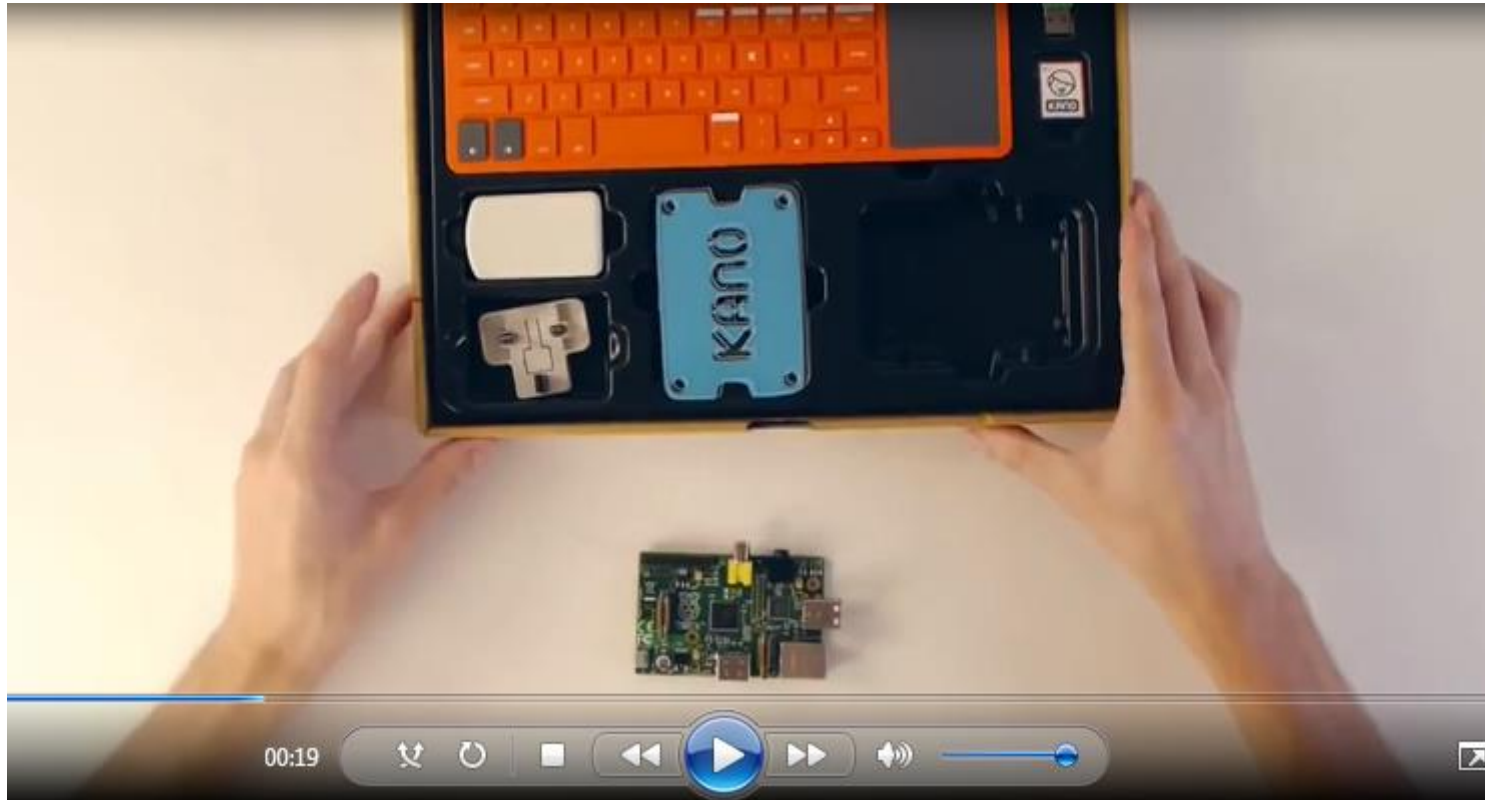
Desempenho dos processadores

- O ganho que tivemos com esse avanço espetacular na tecnologia tem quatro aspectos:
 - Aumento do poder computacional aos usuários. Os processadores top de linha de hoje são melhores que os supercomputadores de 10 anos atrás
 - Gerou novas classes de computadores: workstations em 1980. Depois os laptops. Nos últimos anos os smartphones e tablets. Há também os ultra-minicomputadores para sistemas embarcados (Ex.: Raspberry Pi e Arduino)
 - Domínio dos sistemas baseados em microprocessadores, até mesmo os supercomputadores, reduzindo-se os custos dos sistemas de alta performance
 - Desenvolvimento de software: troca de performance por produtividade com linguagens como Java, C#, Python e Ruby (script)

Do desktop de 1980 aos smartphones



Video: Sistema com Raspberry Pi



How to Make a Computer (in 107 Seconds)

<https://www.youtube.com/watch?v=iNc6NRX2JG4>

Classes de Computadores

- Hoje temos 3 categorias principais de computadores
- Cada categoria se caracteriza por diferentes aplicações, diferentes requisitos, e diferentes tecnologias
- A figura abaixo mostra um sumário das cinco categorias e suas características custo e criticidade.

- Maior crescimento
- Maior gama: custo/desempenho
- Minimização de memória e consumo de potência

Feature	Desktop	Server	Embedded
Price of system	\$500-\$5000	\$5000-\$5,000,000	\$10-\$100,000 (including network routers at the high end)
Price of microprocessor module	\$50-\$500 (per processor)	\$200-\$10,000 (per processor)	\$0.01-\$100 (per processor)
Critical system design issues	Price-performance, graphics performance	Throughput, availability, scalability	Price, power consumption, application-specific performance

Arquitetura do Conjunto de Instruções

- A principal característica de um processador e de sua arquitetura é a arquitetura do conjunto de instruções, ou *Instruction Set Architecture* (ISA).
- O ISA é a interface de programação, e é considerado a fronteira entre o software e o hardware
- Duas classes principais existem hoje:
 - Register-memory – permite o acesso a memória como parte da maioria das instruções. Ex.: 80x86
 - Load-store – o acesso a memória só é feito através das instruções load e store. Ex.: ARM e MIPS
- Endereçamento: a maioria usa endereçamento de byte. MIPS e ARM requerem que os dados estejam alinhados. O 80x86 não requer o alinhamento, mas a velocidade de acesso é menor caso estejam desalinhados

ISA

- Modos de endereçamento: além de especificar registradores e operandos constantes, o modo de endereçamento especifica o endereço de um dado.
- Por exemplo: os modos de endereçamento do MIPS são: Registrador, Imediato e Deslocamento. Neste último uma constante é somada ao registrador para formar o endereço de memória
- Tipos e tamanhos dos operandos: ASCII, 16, 32, 64 bits, IEEE754, etc.
- Categorias de operações: transferência de dados, aritmética, lógica, controle e ponto flutuante.
- Um exemplo de ISA que representa a arquitetura RISC e fácil de ser executada em pipeline é o MIPS, exemplificado a seguir.

Registradores do MIPS

Name	Register Number	Usage	Preserved on call
\$zero	0	the constant value 0	n.a.
\$at	1	reserved for the assembler	n.a.
\$v0-\$v1	2-3	value for results and expressions	no
\$a0-\$a3	4-7	arguments (procedures/functions)	yes
\$t0-\$t7	8-15	temporaries	no
\$s0-\$s7	16-23	saved	yes
\$t8-\$t9	24-25	more temporaries	no
\$k0-\$k1	26-27	reserved for the operating system	n.a.
\$gp	28	global pointer	yes
\$sp	29	stack pointer	yes
\$fp	30	frame pointer	yes
\$ra	31	return address	yes

ISA do MIPS150 (Um MIPS reduzido)

31	26	25	21	20	16	15	11	10	6	5	0	
opcode		rs		rt		rd		shamt		funct		R-type
opcode		rs		rt		immediate						I-type
opcode		target										J-type

Load and Store Instructions

100000	base	dest	signed offset	LB rt, offset(rs)
100001	base	dest	signed offset	LH rt, offset(rs)
100011	base	dest	signed offset	LW rt, offset(rs)
100100	base	dest	signed offset	LBU rt, offset(rs)
100101	base	dest	signed offset	LHU rt, offset(rs)
101000	base	dest	signed offset	SB rt, offset(rs)
101001	base	dest	signed offset	SH rt, offset(rs)
101011	base	dest	signed offset	SW rt, offset(rs)

I-Type Computational Instructions

001001	src	dest	signed immediate	ADDIU rt, rs, signed-imm.
001010	src	dest	signed immediate	SLTI rt, rs, signed-imm.
001011	src	dest	signed immediate	SLTIU rt, rs, signed-imm.
001100	src	dest	zero-ext. immediate	ANDI rt, rs, zero-ext-imm.
001101	src	dest	zero-ext. immediate	ORI rt, rs, zero-ext-imm.
001110	src	dest	zero-ext. immediate	XORI rt, rs, zero-ext-imm.
001111	00000	dest	zero-ext. immediate	LUI rt, zero-ext-imm.

ISA do MIPS150 (Um MIPS reduzido)

R-Type Computational Instructions

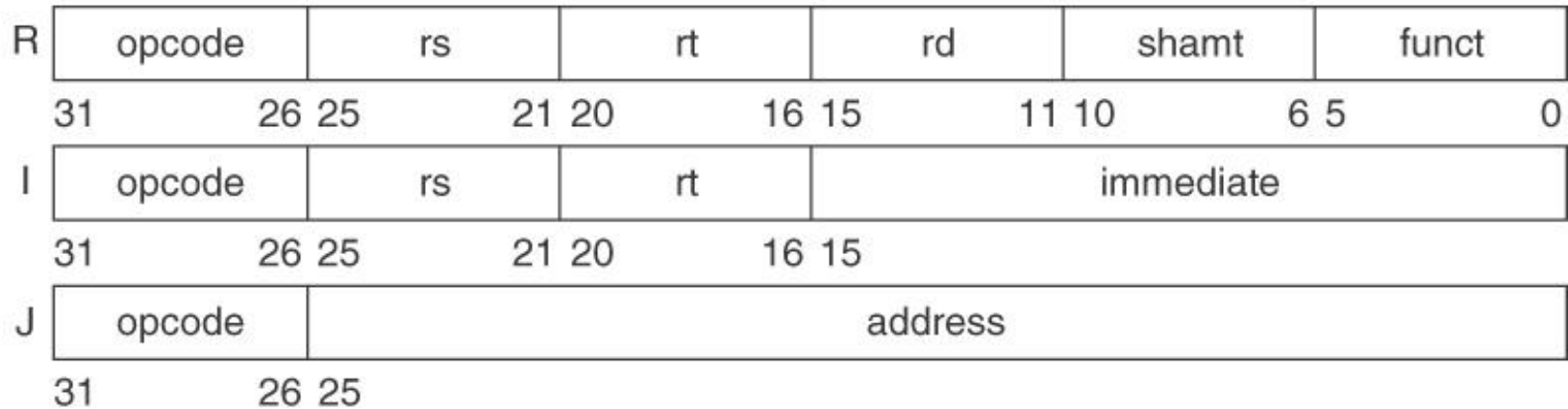
000000	00000	src	dest	shamt	000000	SLL rd, rt, shamt
000000	00000	src	dest	shamt	000010	SRL rd, rt, shamt
000000	00000	src	dest	shamt	000011	SRA rd, rt, shamt
000000	rshamt	src	dest	00000	000100	SLLV rd, rt, rs
000000	rshamt	src	dest	00000	000110	SRLV rd, rt, rs
000000	rshamt	src	dest	00000	000111	SRAV rd, rt, rs
000000	src1	src2	dest	00000	100001	ADDU rd, rs, rt
000000	src1	src2	dest	00000	100011	SUBU rd, rs, rt
000000	src1	src2	dest	00000	100100	AND rd, rs, rt
000000	src1	src2	dest	00000	100101	OR rd, rs, rt
000000	src1	src2	dest	00000	100110	XOR rd, rs, rt
000000	src1	src2	dest	00000	100111	NOR rd, rs, rt
000000	src1	src2	dest	00000	101010	SLT rd, rs, rt
000000	src1	src2	dest	00000	101011	SLTU rd, rs, rt

Jump and Branch Instructions

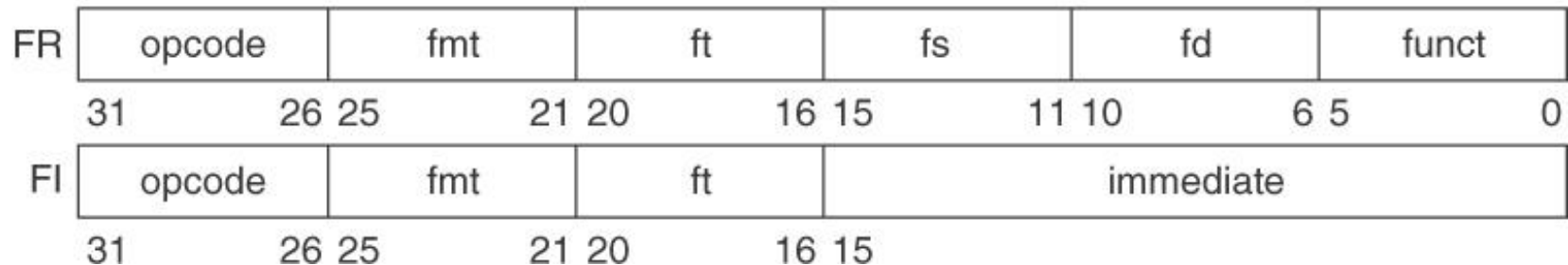
000010	target					J target
000011	target					JAL target
000000	src	00000	00000	00000	001000	JR rs
000000	src	00000	dest	00000	001001	JALR rd, rs
000100	src1	src2	signed offset			BEQ rs, rt, offset
000101	src1	src2	signed offset			BNE rs, rt, offset
000110	src	00000	signed offset			BLEZ rs, offset
000111	src	00000	signed offset			BGTZ rs, offset
000001	src	00000	signed offset			BLTZ rs, offset
000001	src	00001	signed offset			BGEZ rs, offset

Formato das Instruções

Basic instruction formats



Floating-point instruction formats



Exemplo de programa em Assembly do MIPS

```
Loop: g = g + A[i];  
      i = i + j;  
      if (i != h) goto Loop;
```

```
g: $s1  
h: $s2  
i: $s3  
j: $s4  
Base of A: $s5
```

```
Loop: add $t1, $s3, $s3 # $t1 = 2 * i  
      add $t1, $t1, $t1 # $t1 = 4 * i  
      add $t1, $t1, $s5 # $t1 = address of A[i]  
      lw  $t0, 0($t1)  # $t0 = A[i]  
      add $s1, $s1, $t0 # g = g + A[i]  
      add $s3, $s3, $s4 # i = i + j  
      bne $s3, $s2, Loop # go to Loop if i != h
```

Energia e Potência nos Microprocessadores

- Energia Estática x Dinâmica
- A principal é a dinâmica, causada pelo chaveamento dos transistores

$$Energia_{dinâmica} = k.V^2$$

$$Potência_{dinâmica} = k.V^2 freq$$

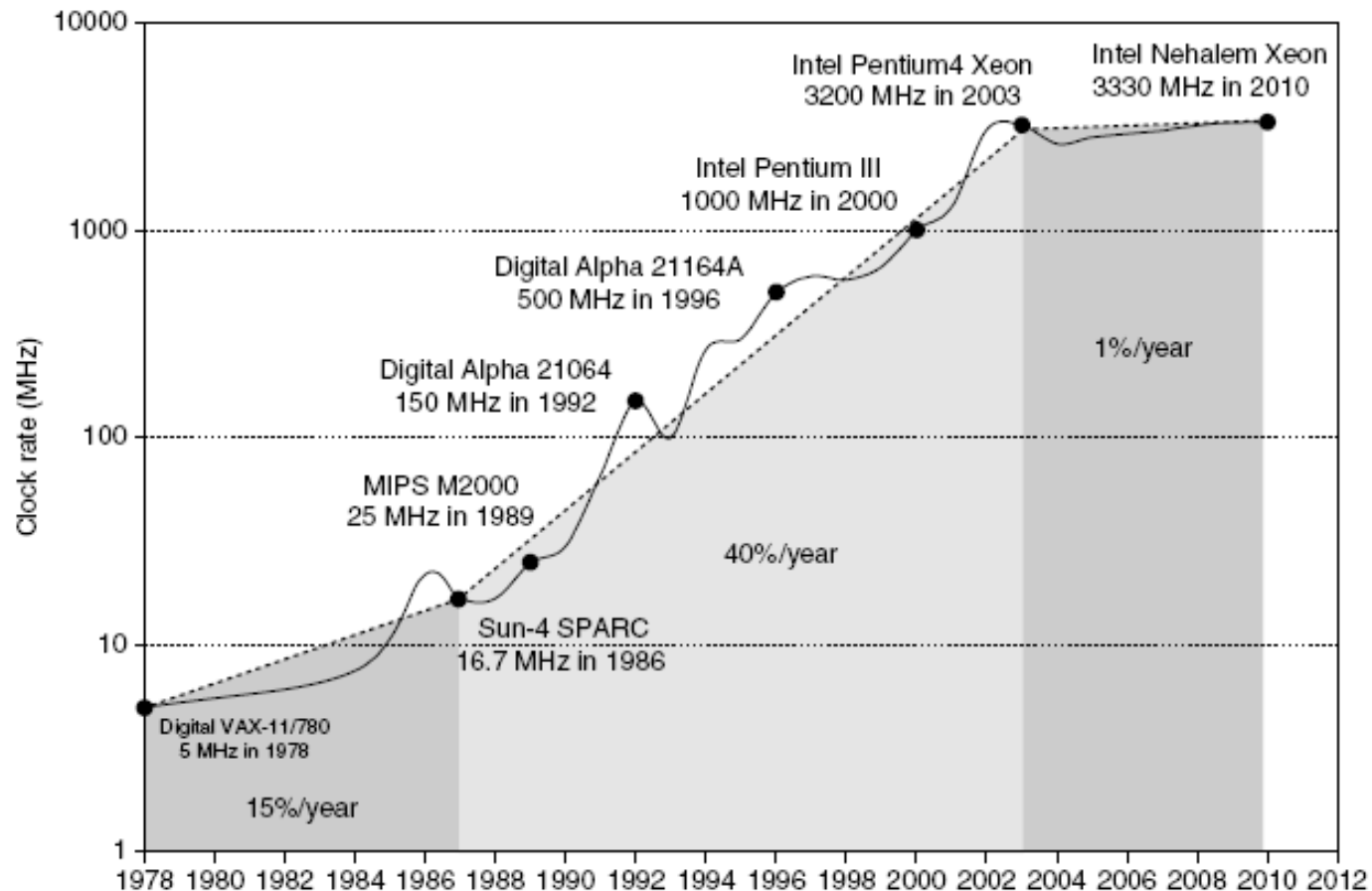
- É possível reduzir a frequência de um processador na mesma proporção quando se reduz a tensão V. Portanto, pode-se aproximar:

$$Potência_{dinâmica} = k.V^3$$

- Como a potência dinâmica depende do cubo da tensão, as tensões de núcleo dos processadores reduziram de 5V para menos de 1V em 20 anos.

Crescimento do Clock.

- A partir de 2004 os uniprocessadores pararam de ser fabricados, dando lugar às arquiteturas *multicore*.



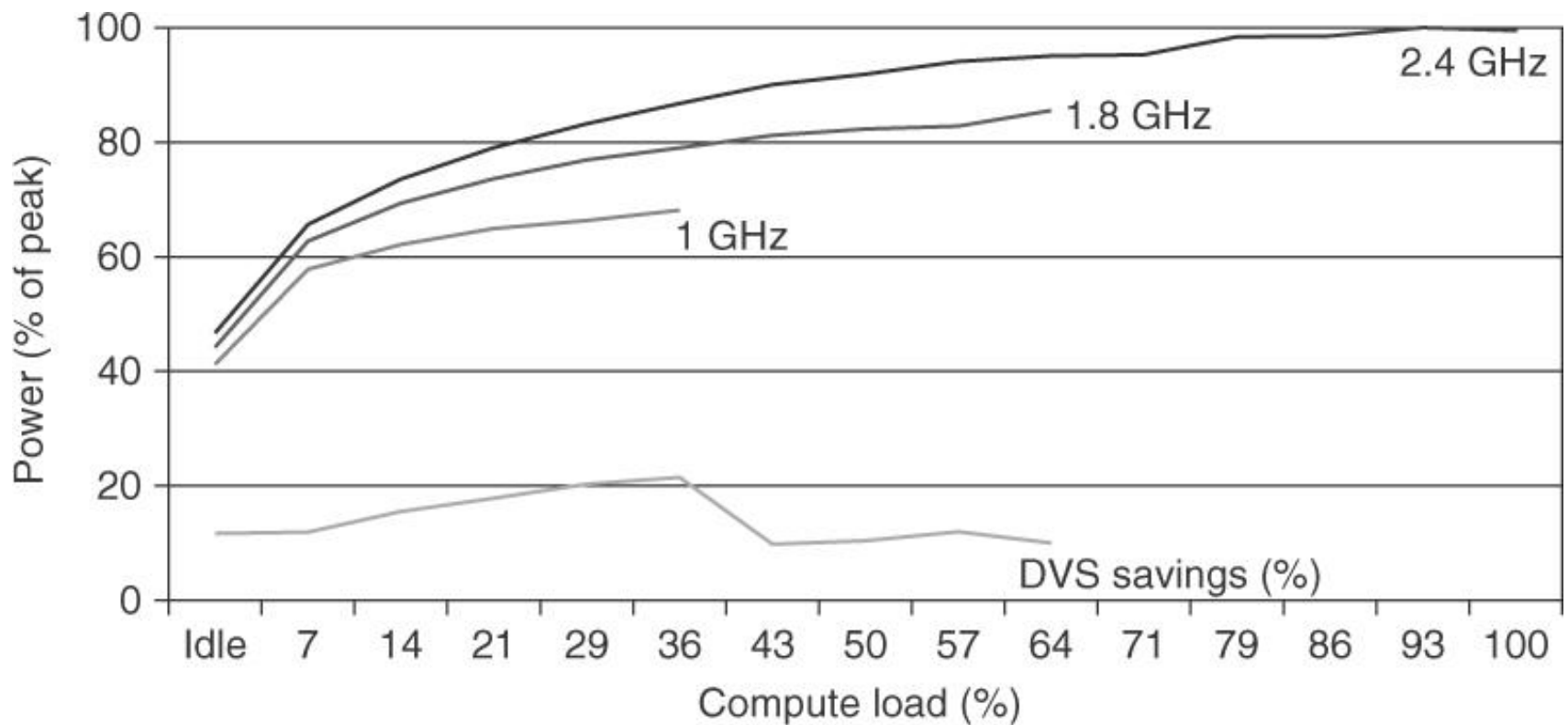
Potência Estática

- A potência estática só é reduzida no projeto do microprocessador e tem sido um alvo das pesquisas atuais nessa área

$$Potência_{estática} = Corrente_{estática} \cdot V$$

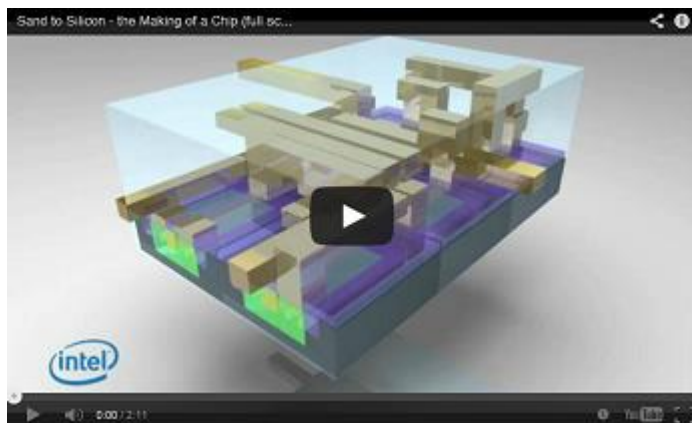
- A corrente estática e a corrente de fuga, que flui mesmo que o transistor esteja em OFF.
- Assim, a potência estática é proporcional ao número de dispositivos (transistores).
- Uma técnica possível para reduzir o consumo de energia incluindo a estática é a chamada *power-gating*, que consiste em se desligar partes do chip quando possível

Consumo de energia de um AMD Opteron com 8GB de DRAM



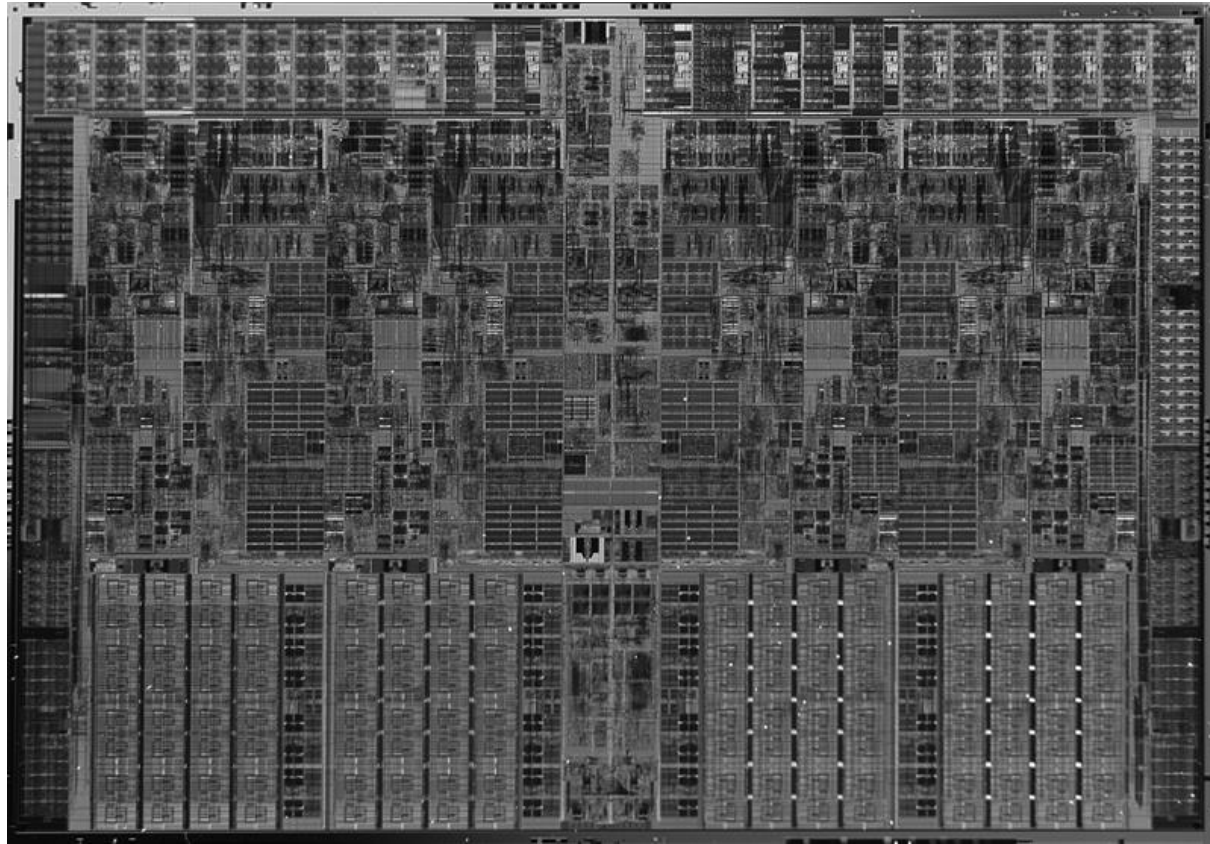
Processo de Fabricação

- Video Intel sobre a fabricação do chip:
- Infográfico



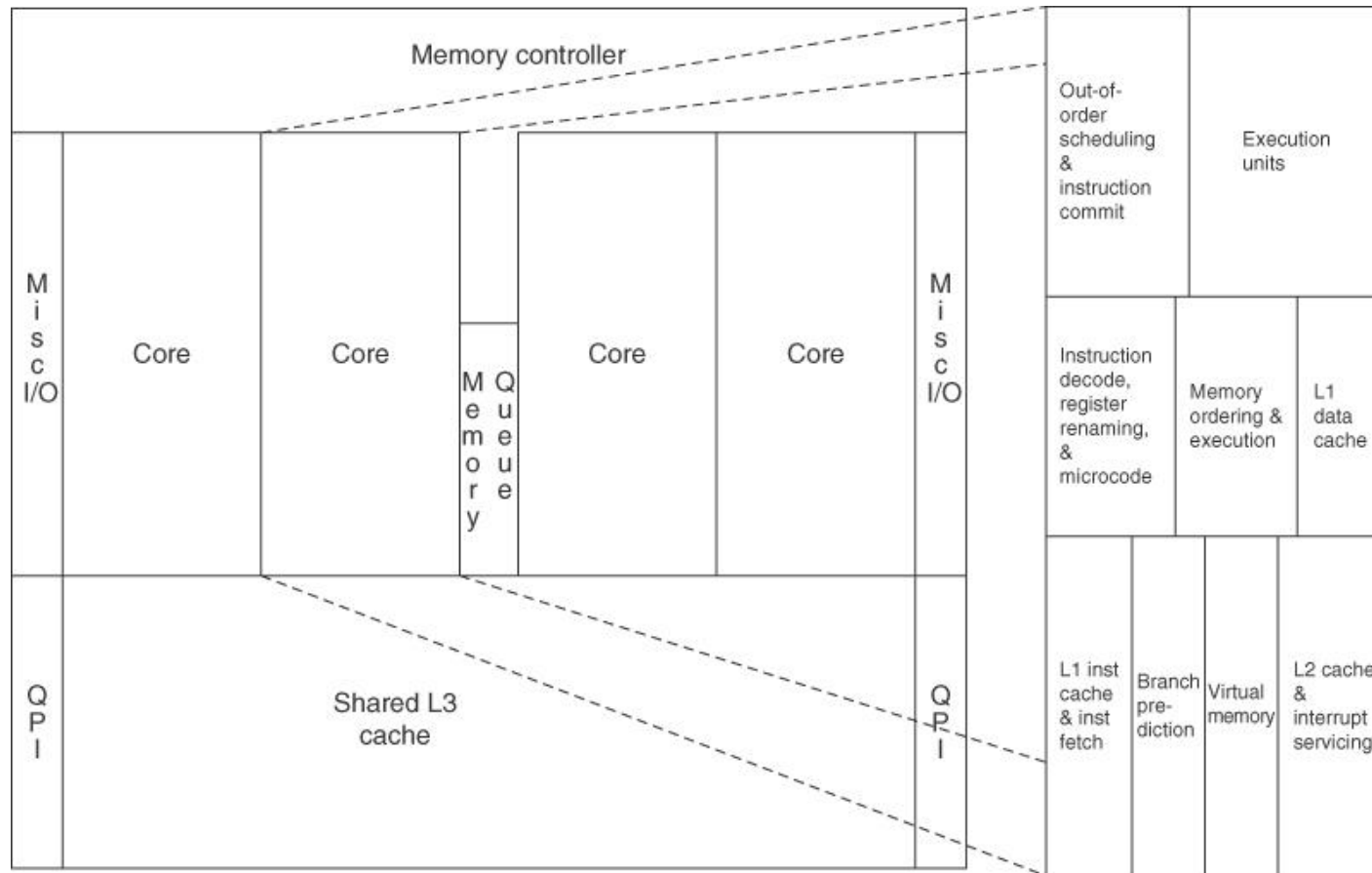
- Fonte: <http://www.tecmundo.com.br/intel/8103-veja-como-sao-produzidos-os-processadores.htm>

Processo de Fabricação



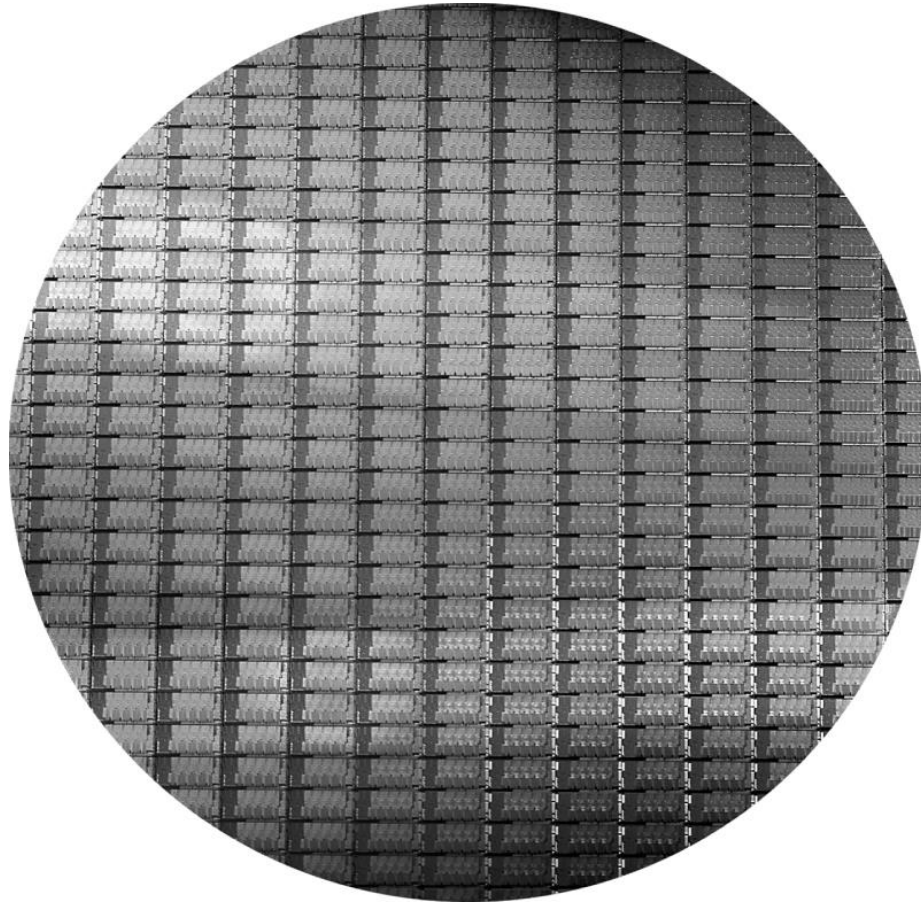
Fotografia da área de chip (die) de um microprocessador Intel Core i7. As dimensões são: 18.9 mm por 13.6 mm (257 mm²) em um processo 45 nm

Processo de Fabricação



- Lay-out do Core i7 da figura anterior, com destaque do lay-out do segundo core ao lado.

Processo de Fabricação



- Wafer de 300mm contendo 280 processadores completos Sandy Bridge (Sucessor da arquitetura Nehalem usado no Core i7), cada um 20,7mm por 10,5mm em um processo 32 nm