



MPSoCs

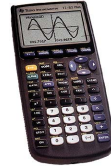
Fernando Gehm Moraes

PUCRS – FACULDADE DE INFORMÁTICA

revisão: 29/abril/2013

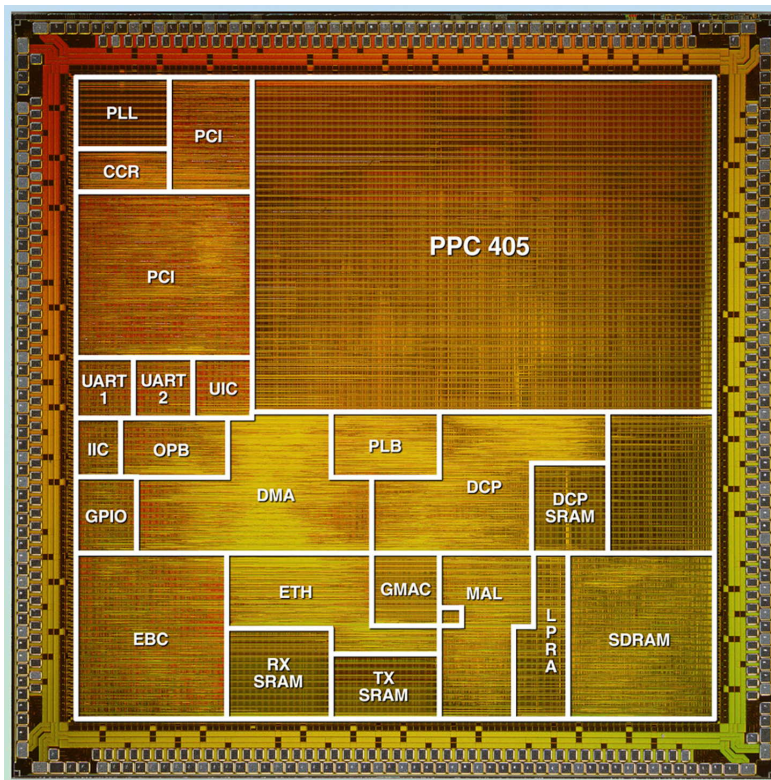


Múltiplas aplicações em um só dispositivo



O que é um SOC (**System-on-a-chip**)?

- Todo sistema integrado em um único circuito integrado (desempenho)
- Projeto modular
- Reuso de IPs

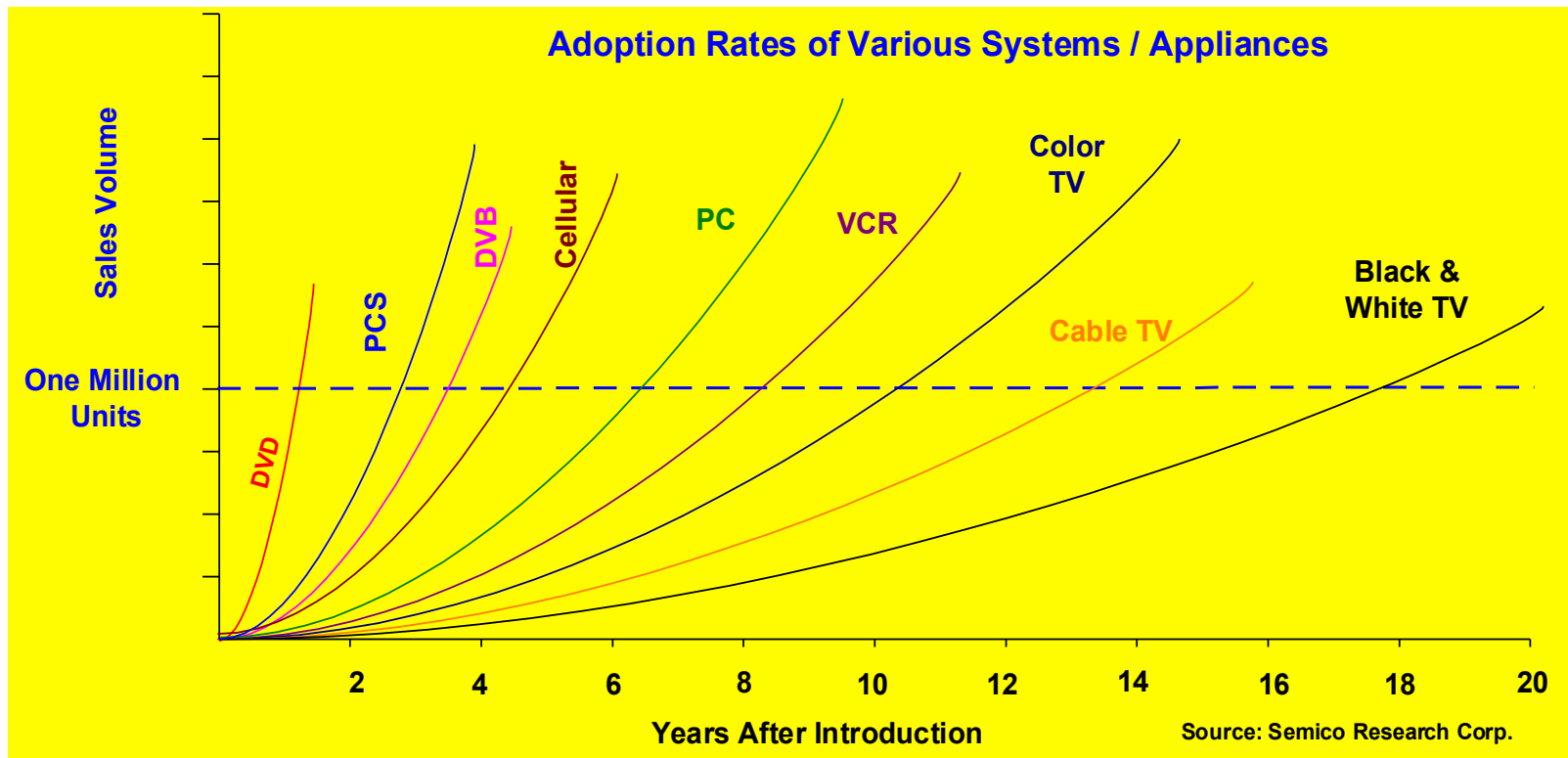


Um SOC pode conter:

- Módulos de propriedade intelectual reutilizáveis
- Processador(es) embarcado(s)
- Memória embarcada
- Software
- Interfaces com o mundo externo (USB, PCI, Ethernet)
- Blocos analógicos
- Hardware programável (FPGAs)

Porque SoCs?

- Time-to-market!



E o IPAD? 28 dias para vender 1 milhão de unidades!

Dedicated hw
1 algorithm



The die photo shows the PPC 405 microprocessor die with the following labeled blocks:

- PLL
- CCR
- PCI
- PCI
- PPC 405
- UART 1
- UART 2
- UIC
- IC
- OPB
- GPIO
- DMA
- PLB
- DCP
- DCP SRAM
- EBC
- ETH
- GMAC
- MAL
- LP RAM
- SRAM
- RX SRAM
- TX SRAM

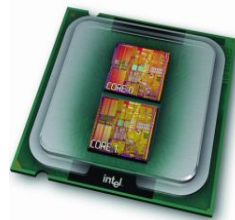
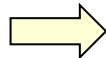
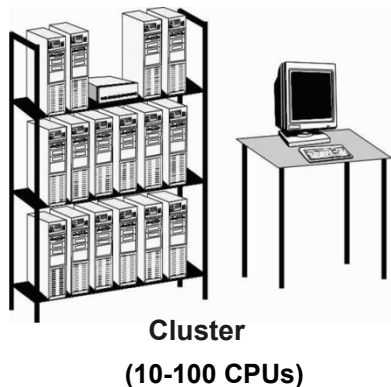
80's

Platform design
Target multiple applications

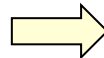


MPSoC – Multiprocessor system-on-chip

- Sistema multiprocessado integrado como um SoC
- Elementos de processamento + IPs + Infra-estrutura de Comunicação
- Evolução de *Clusters* (+ recursos = + desempenho)



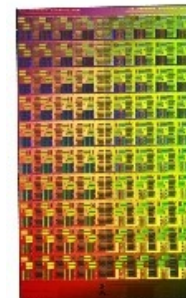
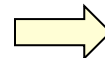
Core Duo Technology
(2-4 CPUs)



Today



MPSoC
(10-100 CPUs)



Chip da Intel
65nm
80-Cores
conectados via NoC
1.28 TFLOPS @ 4GHz
275 mm²

Previsão da evolução dos MPSoCs

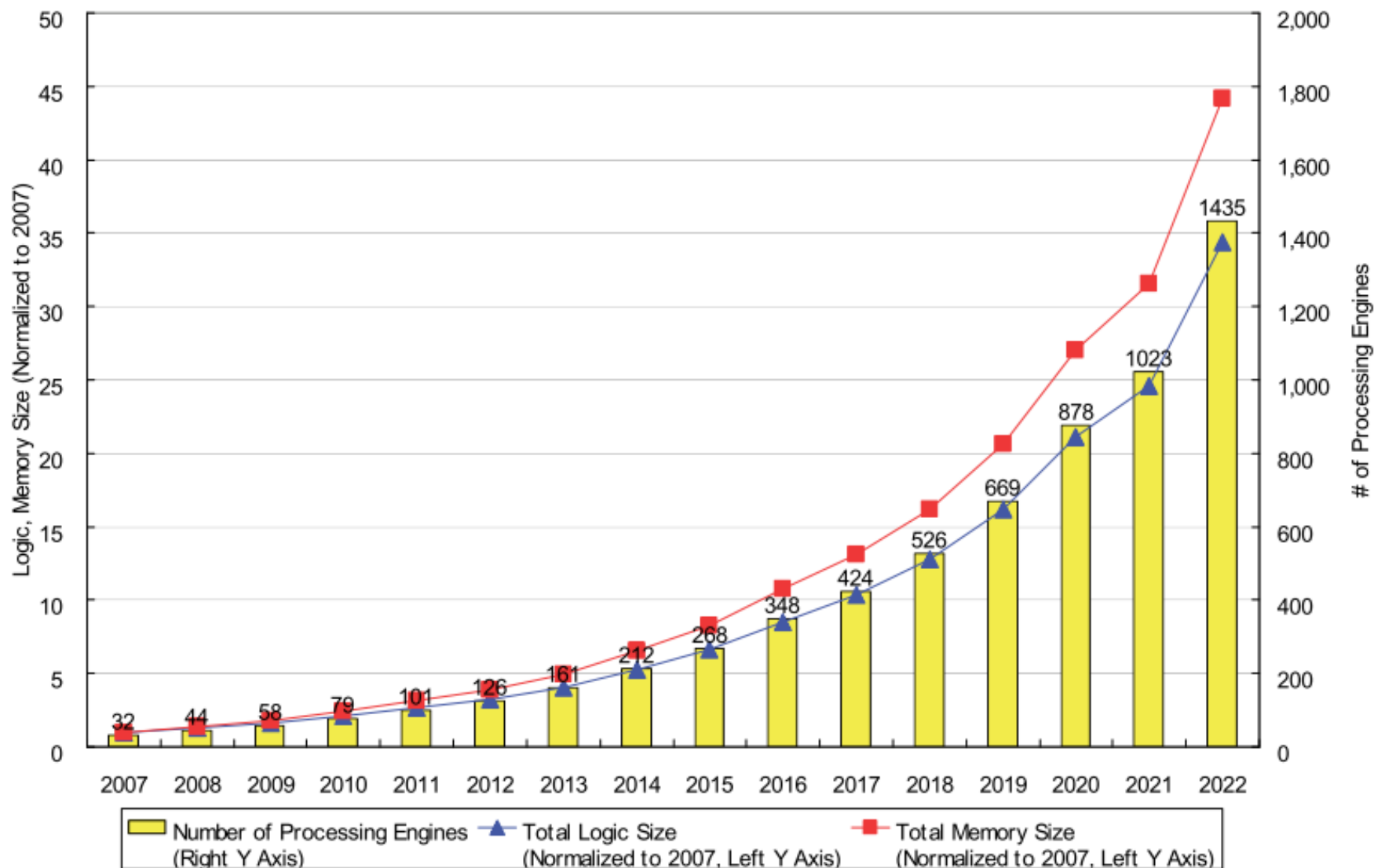


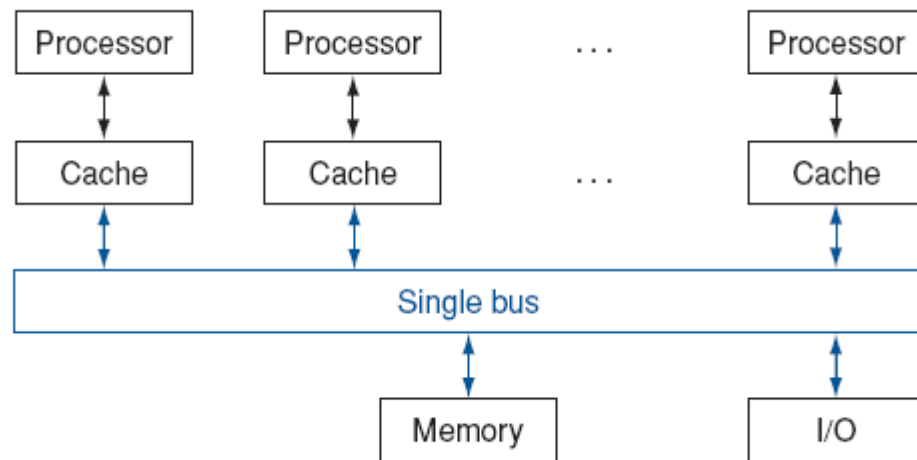
Figure SYSD5 SOC Consumer Portable Design Complexity Trends

Organizações básicas de MPSoCs

- A forma como a comunicação entre os processadores é realizada **define a organização** do MPSoC
- Modelo de comunicação com a memória
 - Memória Unificada – UMA / NUMA
 - Troca de mensagens
- Modelo físico de comunicação
 - Barramento
 - Rede intra-chip, ou NoC (Network-on-chip)

MPSoC conectado por barramento

- Número típico de processadores: 2-32
- Espaço único de endereçamento
- Coerência de *cache* é simples (protocolos *snoop*)

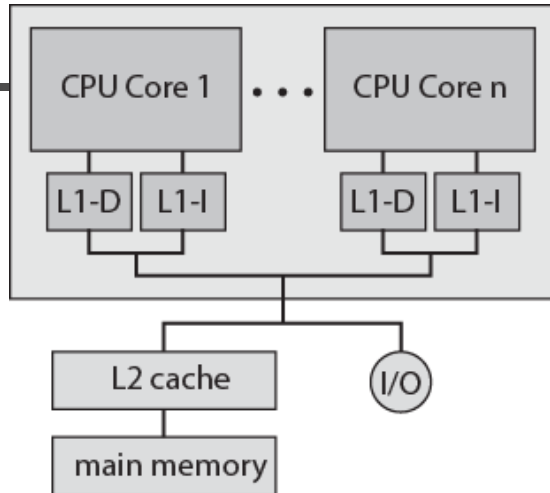


cache coherency Consistency in the value of data between the versions in the caches of several processors.

Programação mais simples, mas não escalável

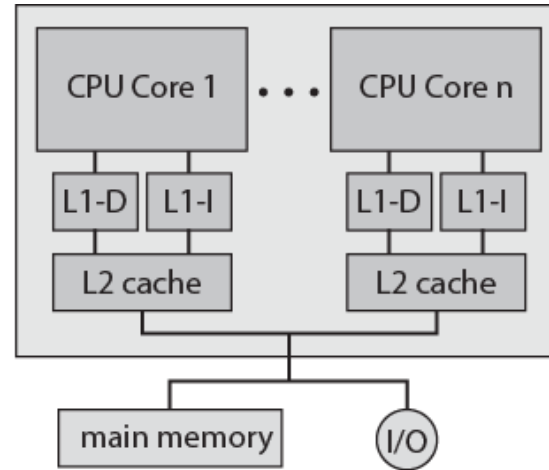
Barramentos e hierarquia de memória

ARM11 MPCore



(a) Dedicated L1 cache

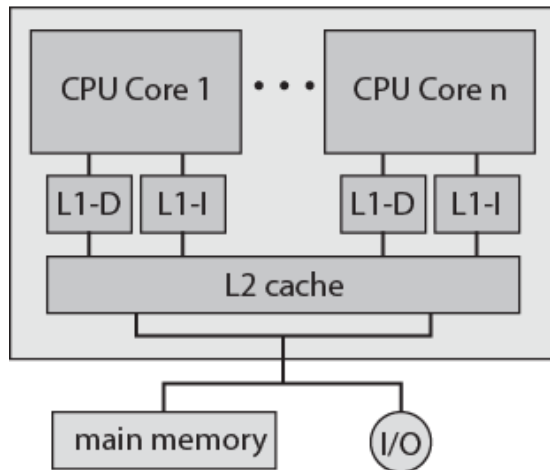
AMD Opteron



(b) Dedicated L2 cache

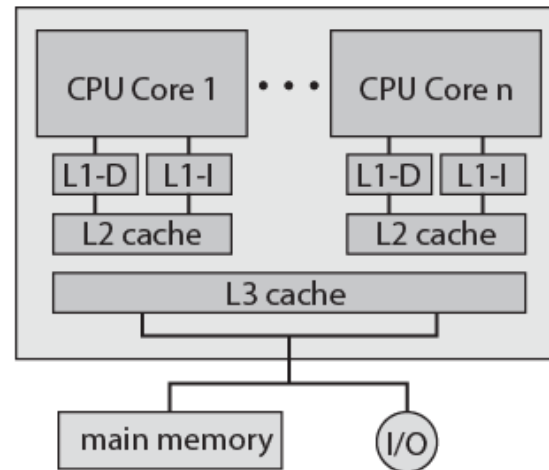
No shared

Intel Core Duo



(c) Shared L2 cache

Intel Core i7

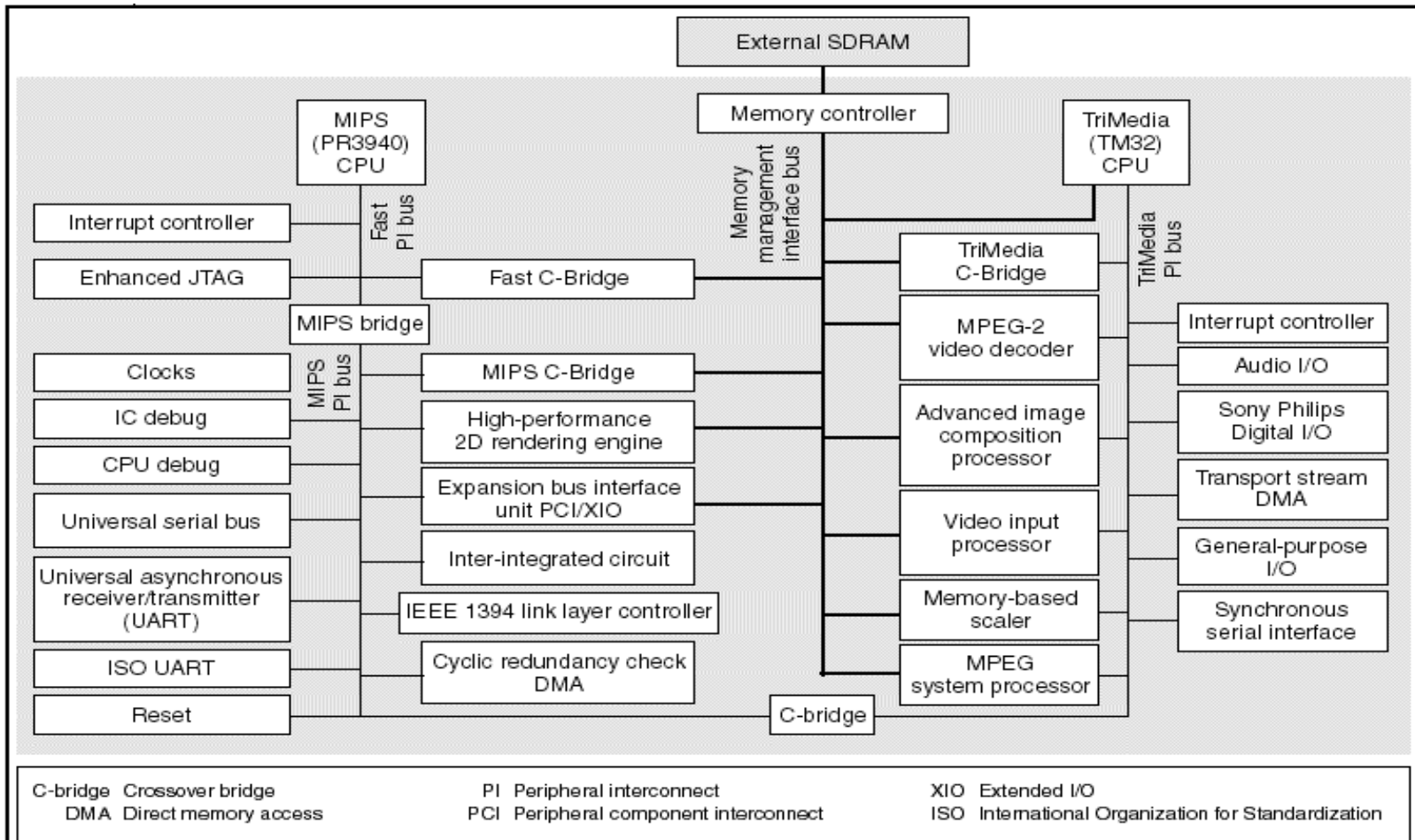


(d) Shared L3 cache

Shared

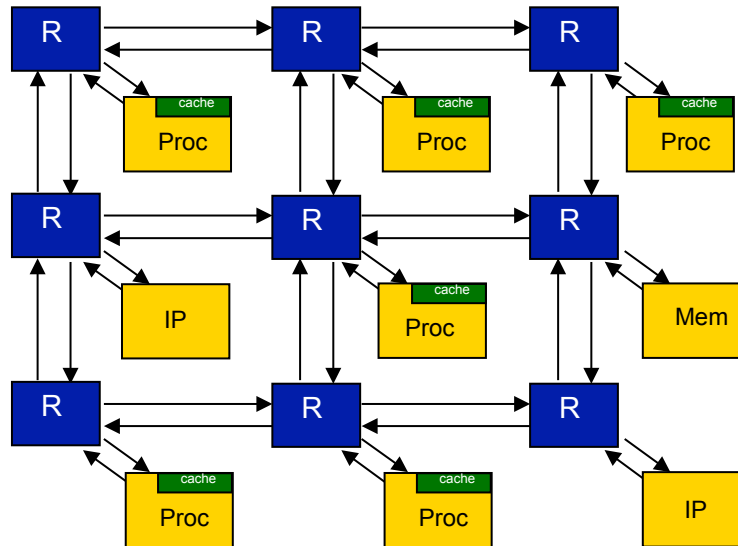
Barramento

- Complexidade crescente da interconexão por barramento



MPSoC conectado por rede (NoC)

- Arranjo de roteadores conforme uma *topologia*
 - Abaixo uma rede malha (*mesh*)
- Elementos de processamento, memórias, IPs conectados aos roteadores (redes diretas)
- Comunicação por troca de mensagens



Programação mais complexa, porém escalável

Organizações básicas de MPSoCs

Category	Choice		Number of processors
Communication model	Message passing		8–2048
	Shared address	NUMA	8–256
		UMA	2–64
Physical connection	Network		8–256
	Bus		2–36

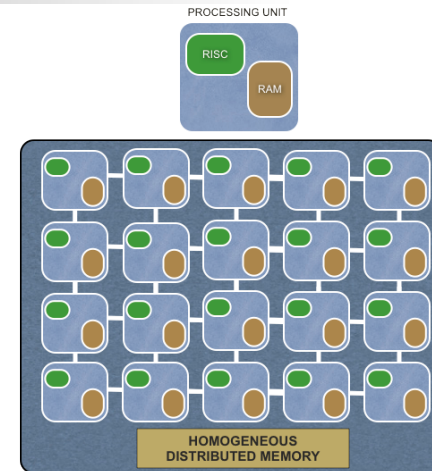
Logo:

- 1) Utilizar comunicação por troca de mensagens
- 2) Utilizar redes intra-chip (NoCs) ao invés de barramento

MPSoC – tipo de processamento

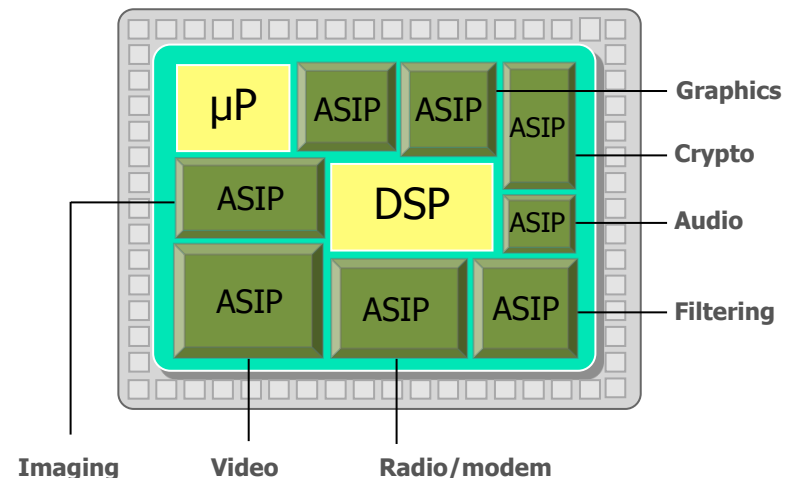
■ Homogêneo

- elementos de processamento (PEs) idênticos
- Pros
 - tarefa de programação – mesma aplicação pode executar em qualquer PE
 - mapeamento das aplicações no MPSoC
- Cons
 - tarefas mais intensivas podem não ter seus requisitos de desempenho atendidos



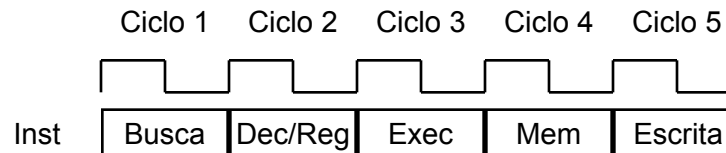
■ Heterogêneo

- PEs distintos (GPPs, DSPs, NPUs)
- Pros
 - PEs dedicados para tarefas específicas
- Cons
 - gerência do sistema
 - programação

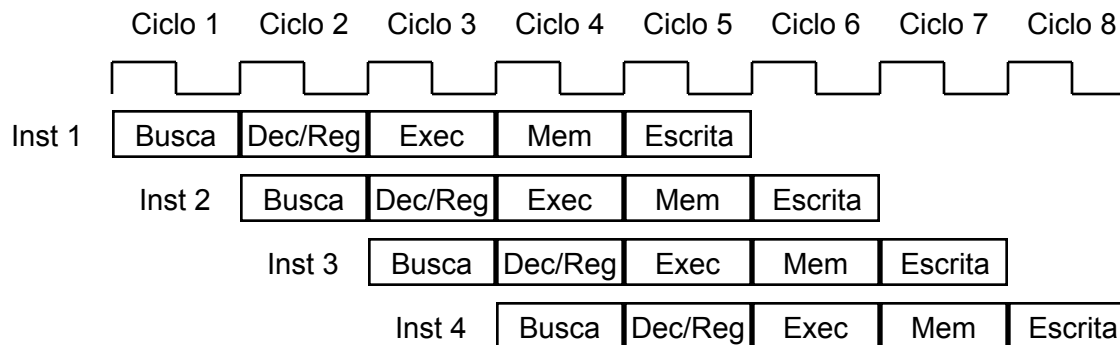


Princípios básicos de paralelismo

CPU Multiciclo



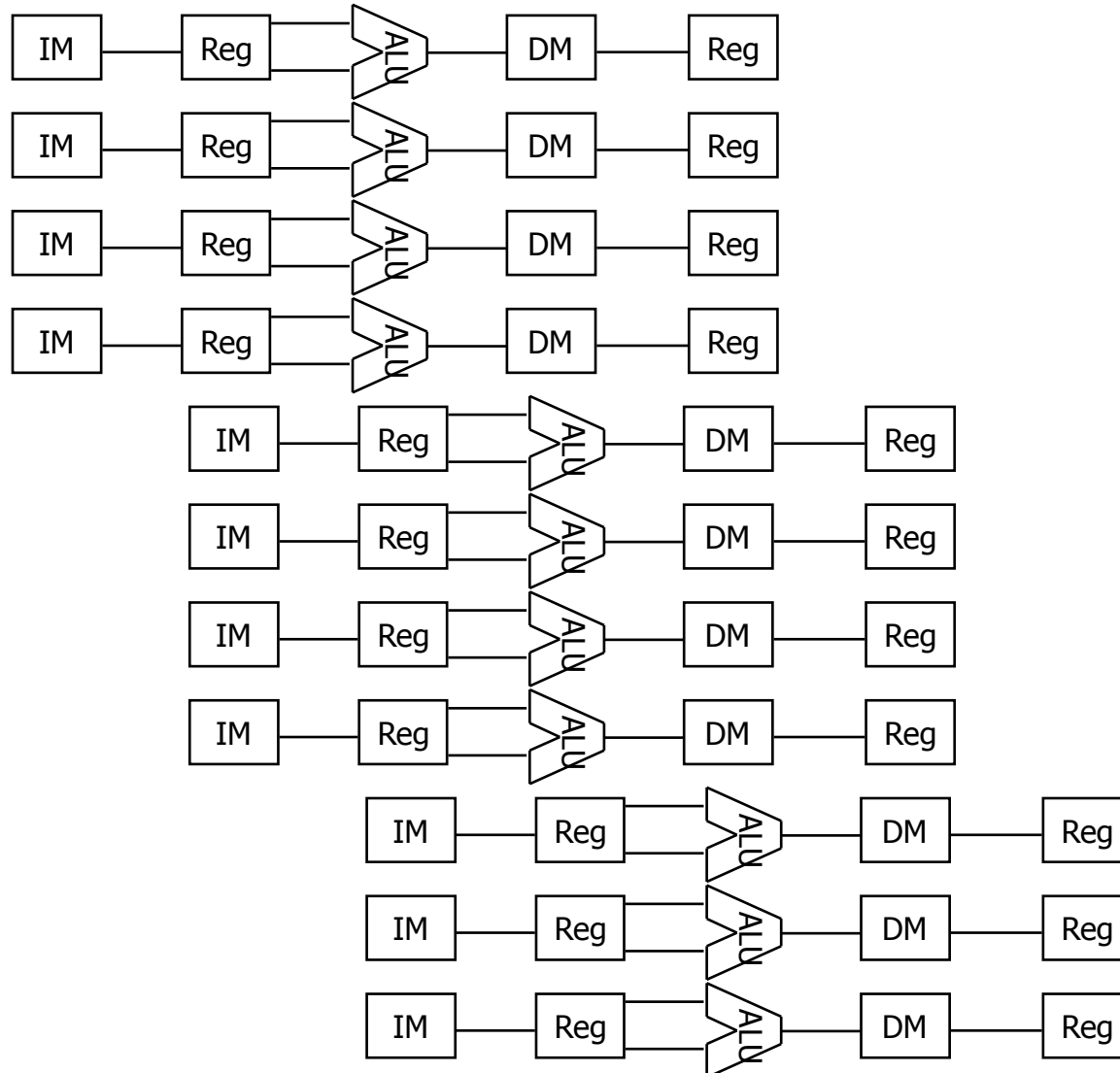
CPU com pipeline



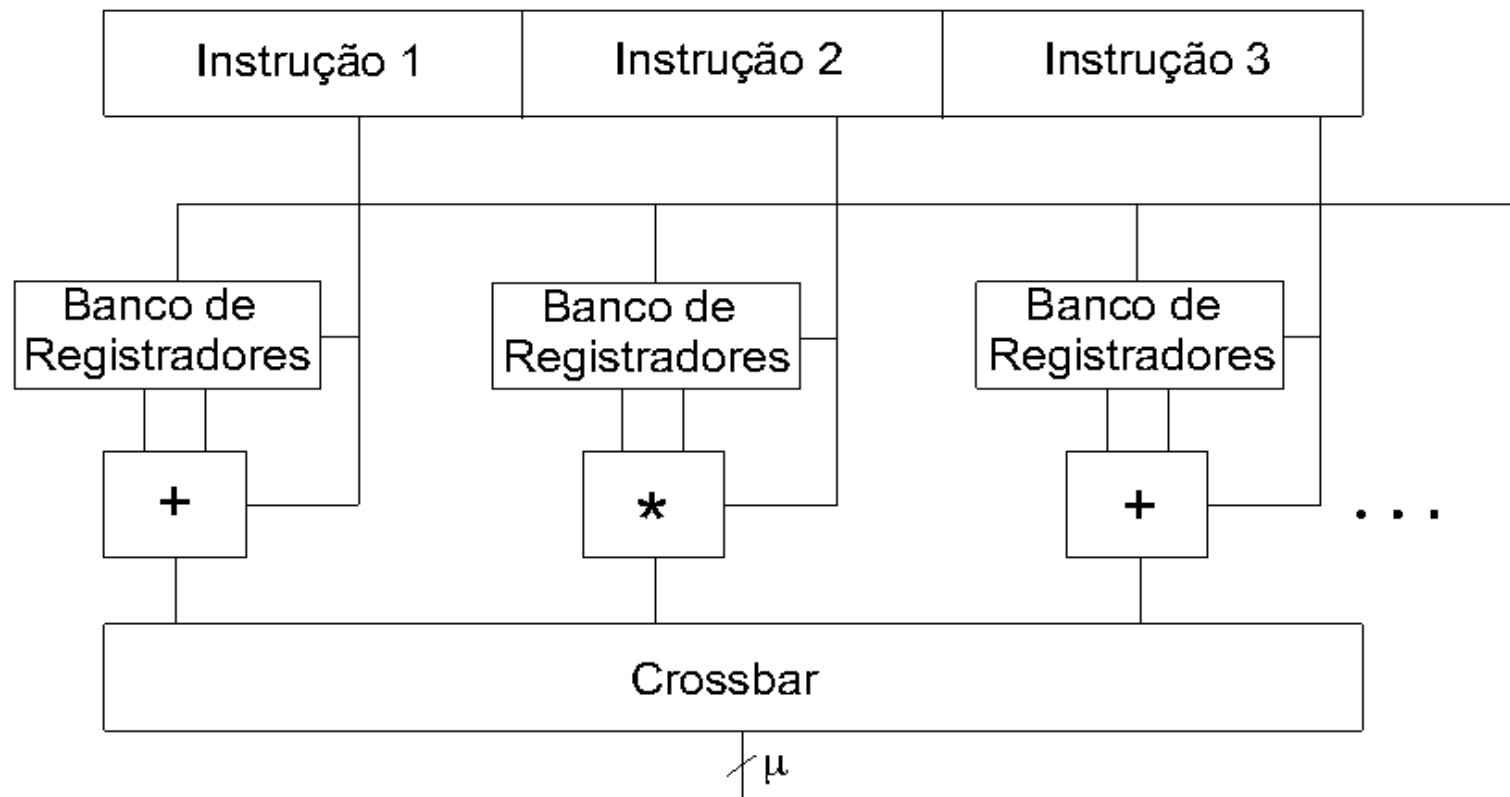
Execução em paralelo!



Execução Superscalar



Máquina VLIW



Como processadores ficaram mais rápidos?

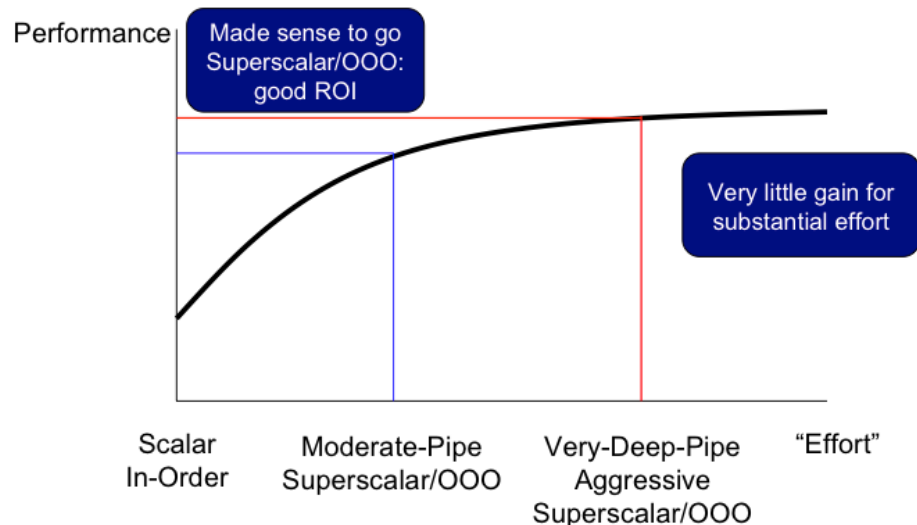
- superpipelining
- superscalar execution
- dynamic scheduling
- multilevel memory caching
- aggressive speculation
- DSP specialization
- fabrication technology

“1º MURO” - ILP WALL

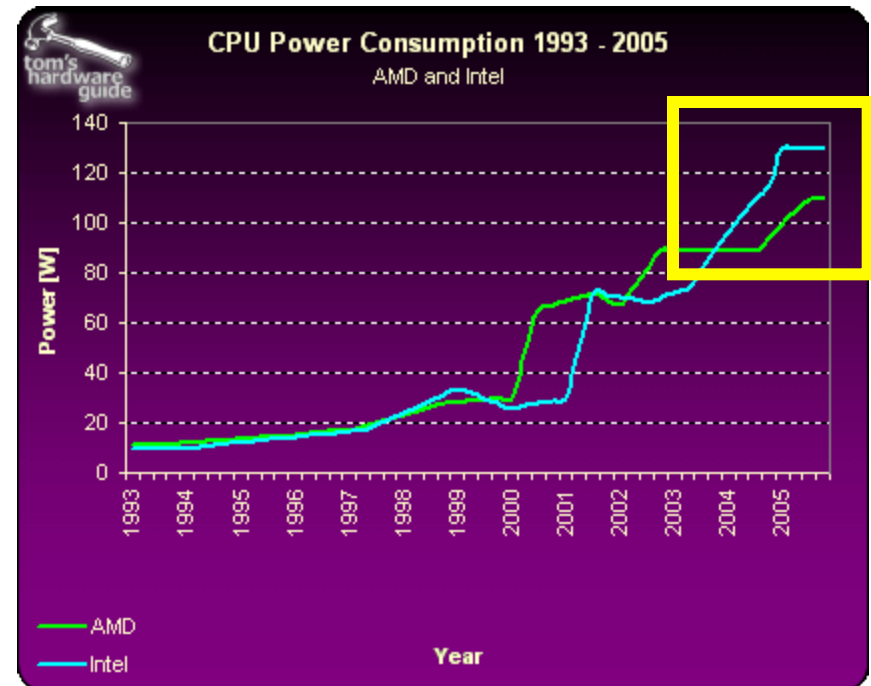
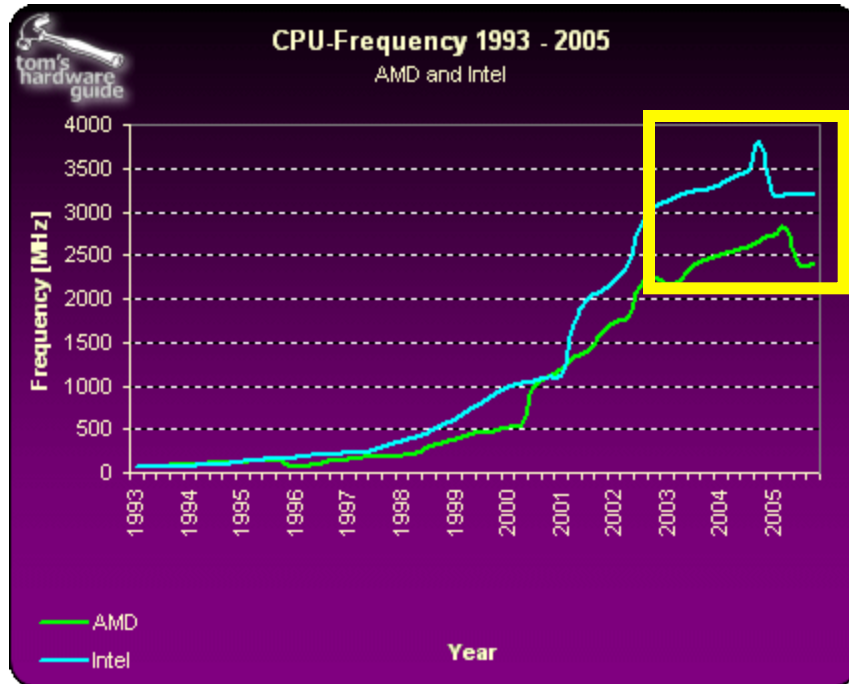
■ Influência do compilador!

- descobrir o paralelismo em máquinas superescalar e VLIW
- a ordem das instruções (dependência) afeta o desempenho

■ “ILP wall”



“2°/3° MUROs”: FREQ. WALL / POWER WALL

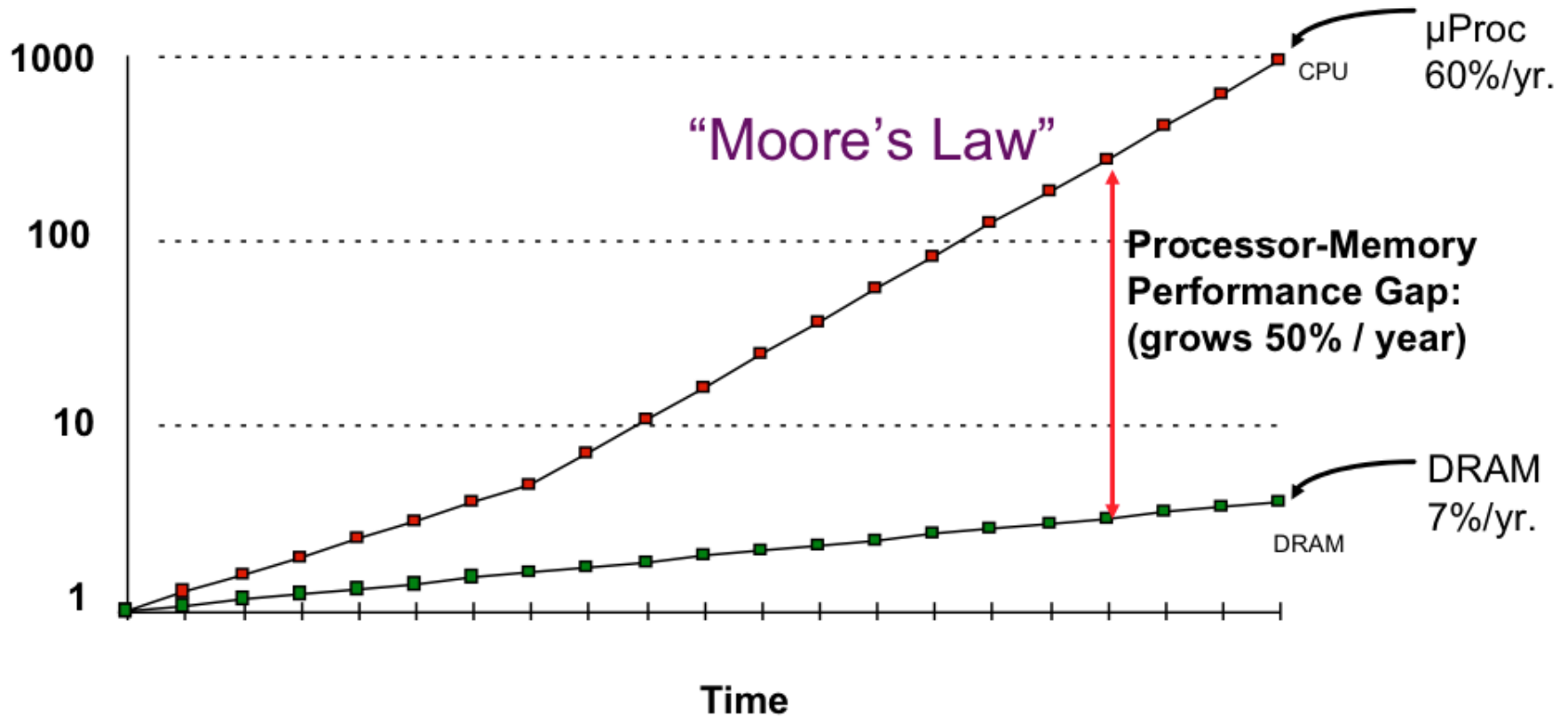


“4º MURO” - MEMORY WALL

	Capacidade	Velocidade
Lógica:	2x em 3 anos	2x em 3 anos
DRAM:	4x em 3 anos	2x em 10 anos
Disco:	4x em 3 anos	2x em 10 anos

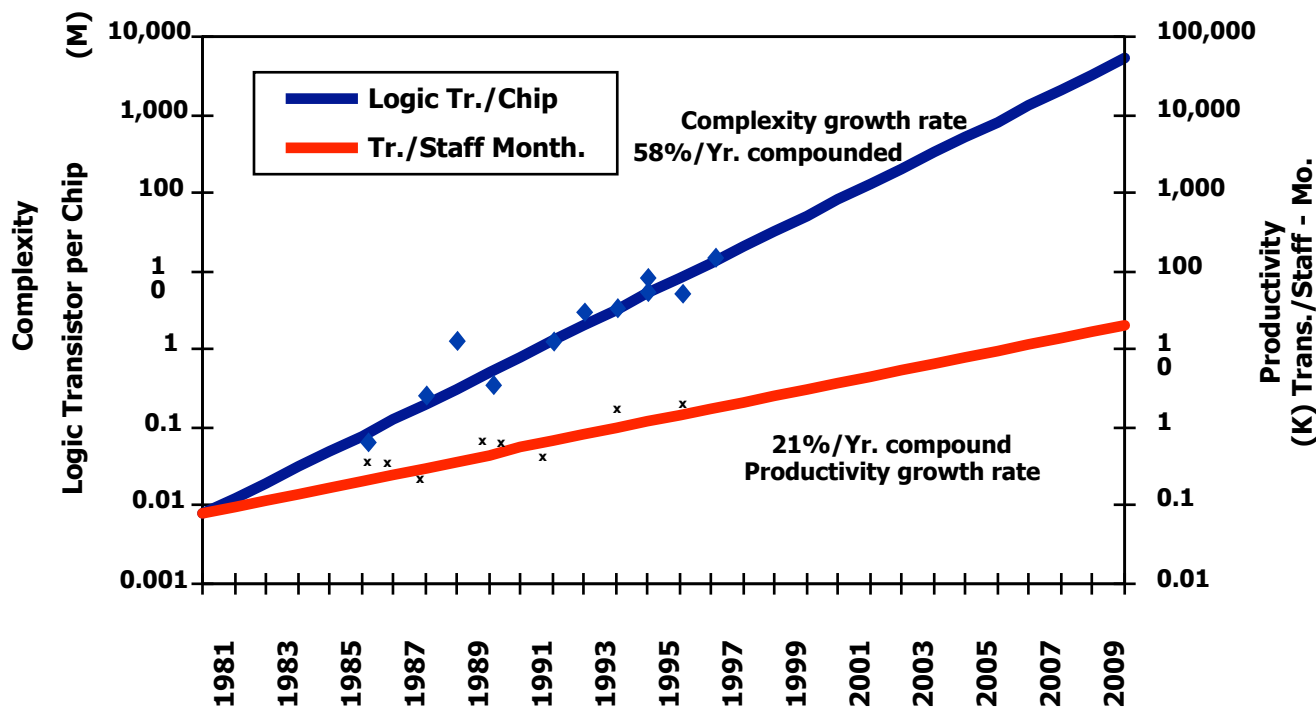
- Fato: memórias grandes são lentas, memórias rápidas são pequenas
- Como criar uma memória grande e rápida (pelo menos na maior parte do tempo)?
 - Hierarquia
 - Paralelismo

“4° MURO” - MEMORY WALL



Mais um problema: *Design Gap*

- Capacidade de integração ultrapassa a capacidade de projeto



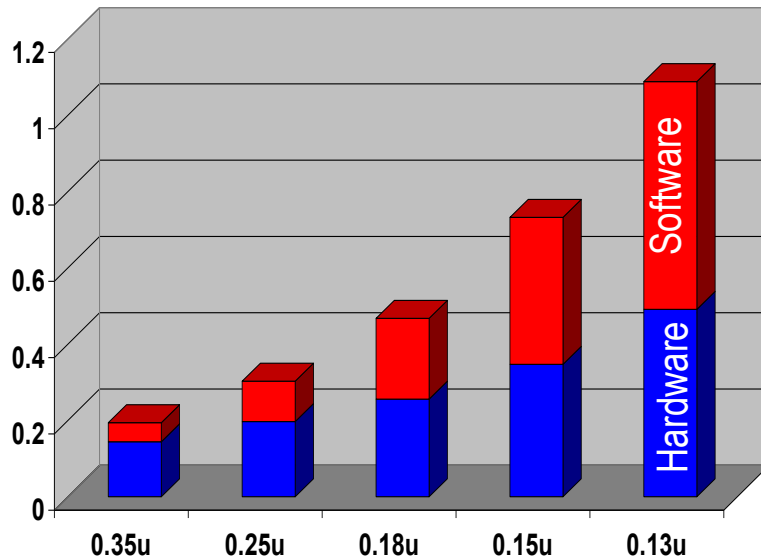
Source: Sematech

Complexity outpaces design productivity

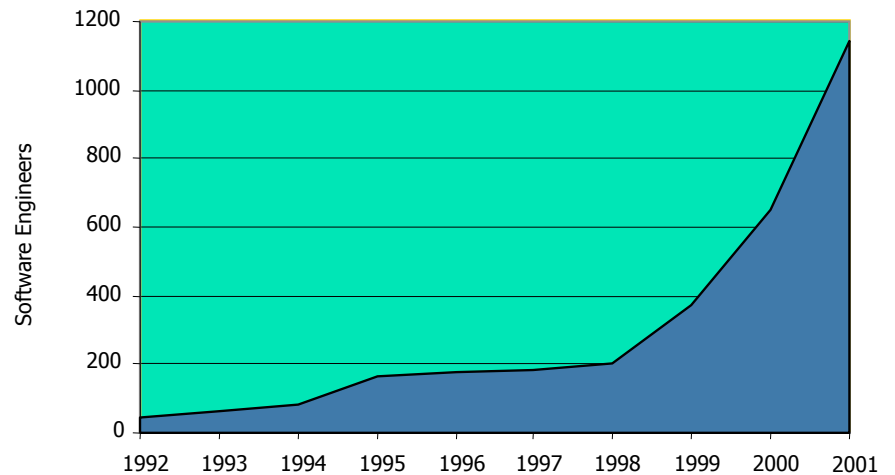
Courtesy, ITRS Roadmap

MORE SOFTWARE THAN HARDWARE

IC Hardware & Software Effort



Growth in SW Engineers

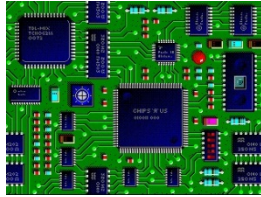


Source: Top Ten Semiconductor Supplier

Então, porque utilizar MPSoCs ?

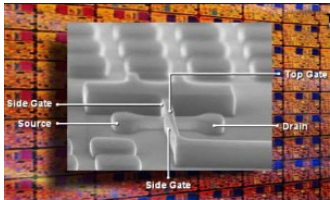
- Quebram o “ILP wall”
 - Múltiplas threads/tarefas executam simultaneamente
 - Paralelismo de *grão-grande* (não mais no nível de instrução)
- Quebram o “frequency wall” e o “power wall”
 - Múltiplos PEs mais lentos e mais simples
 - Escalabilidade obtida pelo aumento do número de PEs e não pelo aumento da frequência
 - Uso de processamento heterogêneo pra aumento de desempenho/redução de potência
- Quebram o “memory wall”
 - Hierarquia de memórias distribuídas e adaptadas às aplicações
 - Replicação do sistema operacional nos PEs
- Diminuem o “design gap”
 - Replicação dos PEs diminui os custos de projeto de verificação
 - Aumenta a importância da rede de comunicação

Receita de MPSoC ☺



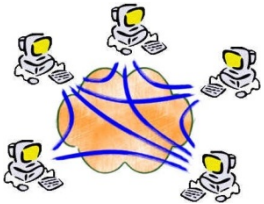
Concepção de circuitos

- Reuso de módulos
- Comunicação



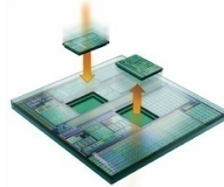
Tecnologia do Silício

- Transistores menores
- Fios mais próximos



Redes de computadores

- Topologias
- Algoritmos de roteamento
- etc



Conceito de SoC

- Tudo no mesmo CI
- Melhor desempenho
- Menor *time-to-market*

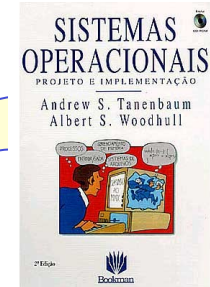


Conceito de cluster

- Replica recurso = maior desempenho
- Balanceamento de carga
- Migração de tarefas



Multidão de Engenheiros

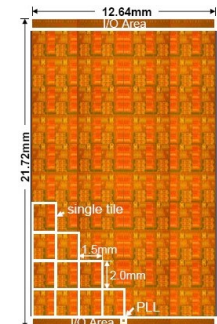


Sistemas Operacionais

- Gerência de recursos
- Escalonamento de processos
- **Mapeamento** de processos



MPSoC



Reusando conceitos antigos para novos cenários

Exemplos industriais de MPSoCs (1/4)

Cell Processor

■ Processador central

■ Power Processing Element

- IBM PowerPC
- Multithreaded (2 Threads)
- Cache L1 (32kb I e 32 Kb D)
- Cache L2- 512 Kb
- Frequência: 3.2 GHz
- Atua como controlador dos processadores periféricos

■ Oito processadores periféricos

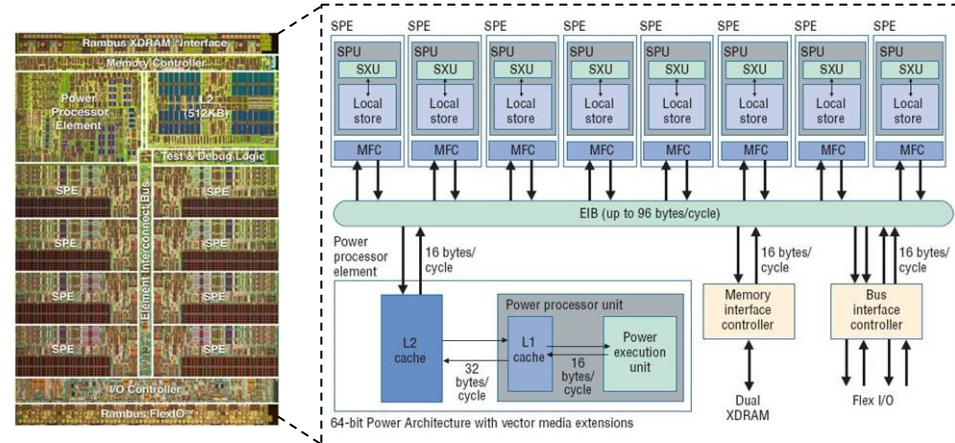
■ Synergistic Processing Elements (SPE)

- Baseados em Processamento Vetorial (SIMD)
- 256 Kb Local Storage
- Controlados por software

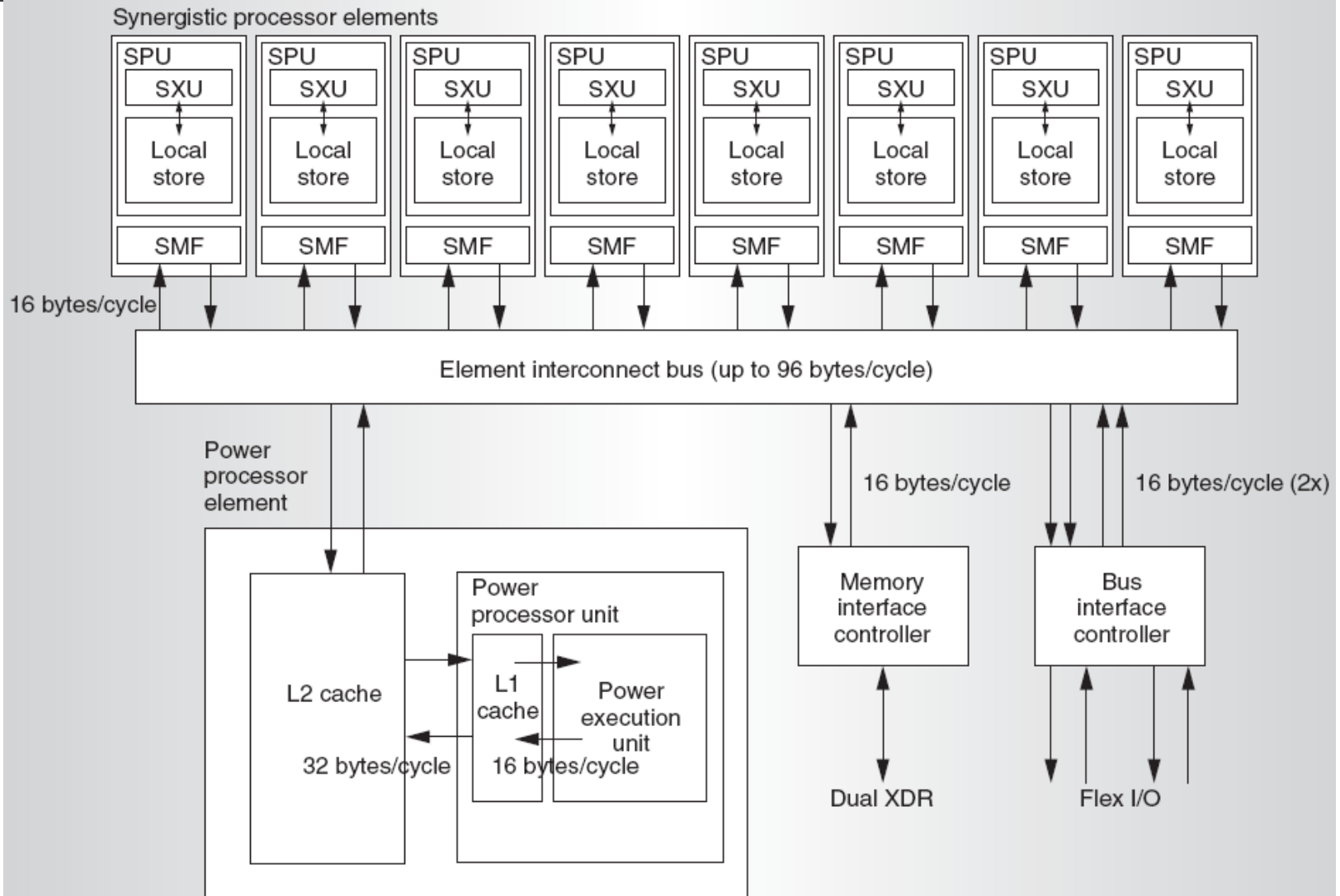
■ Dificuldades para codificar software

- Desenvolvedores de SW são responsáveis pelo gerenciamento dos dados das Local Storages

MPSoC proposto pela IBM para o PlayStation3 (2006)



CELL Processor



Fonte: IEEE Micro Março/Abril 2013

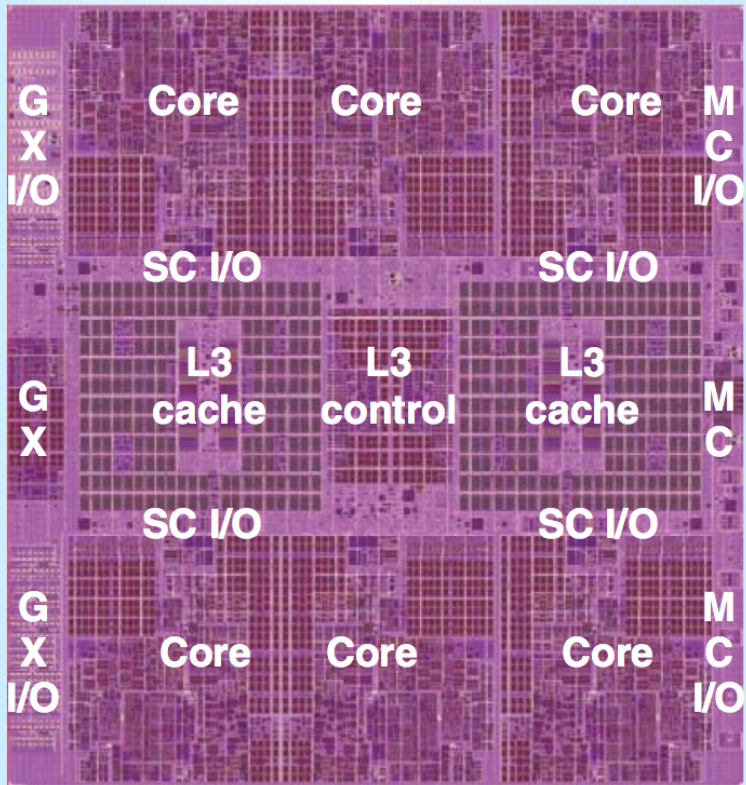


Figure 1. The zEnterprise EC12 (zEC12) microprocessor chip with six processor cores, 48 Mbytes of level-3 (L3) cache, and logic and interfaces connecting to the rest of the system. The chip was fabricated using 32-nm silicon-on-insulator (SOI) technology with roughly 2.75 billion transistors.

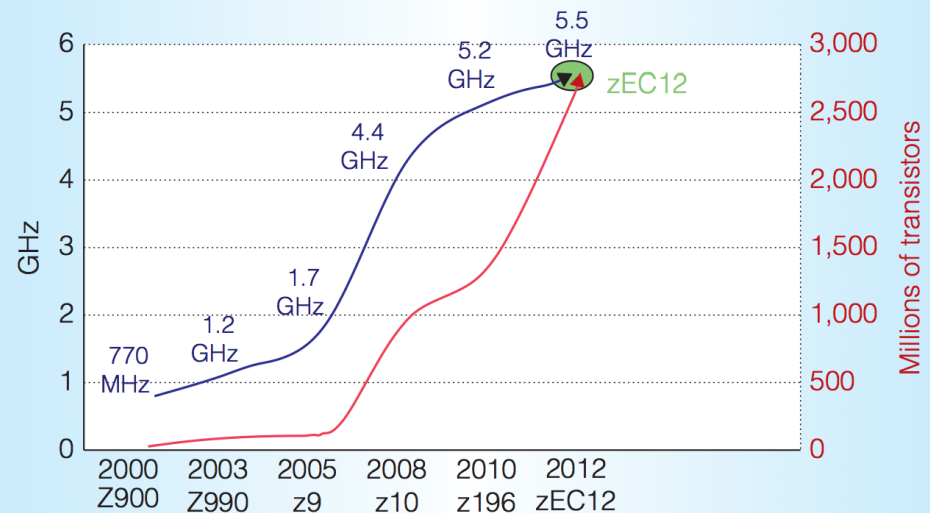
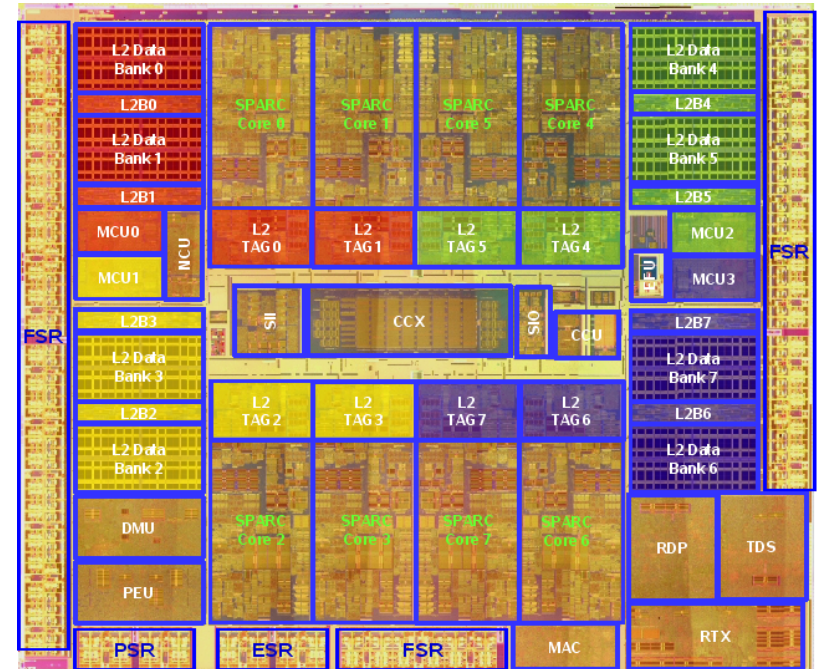


Figure 2. Frequency and transistor counts for the last six generations of microprocessor chips used in IBM's mainframes. System z10 was designed with a deep pipeline microarchitecture to operate at an ultra-high frequency, allowing the next two generations, z196 and zEC12, to continue their frequency improvements.

Exemplos industriais de MPSoCs (2/4)

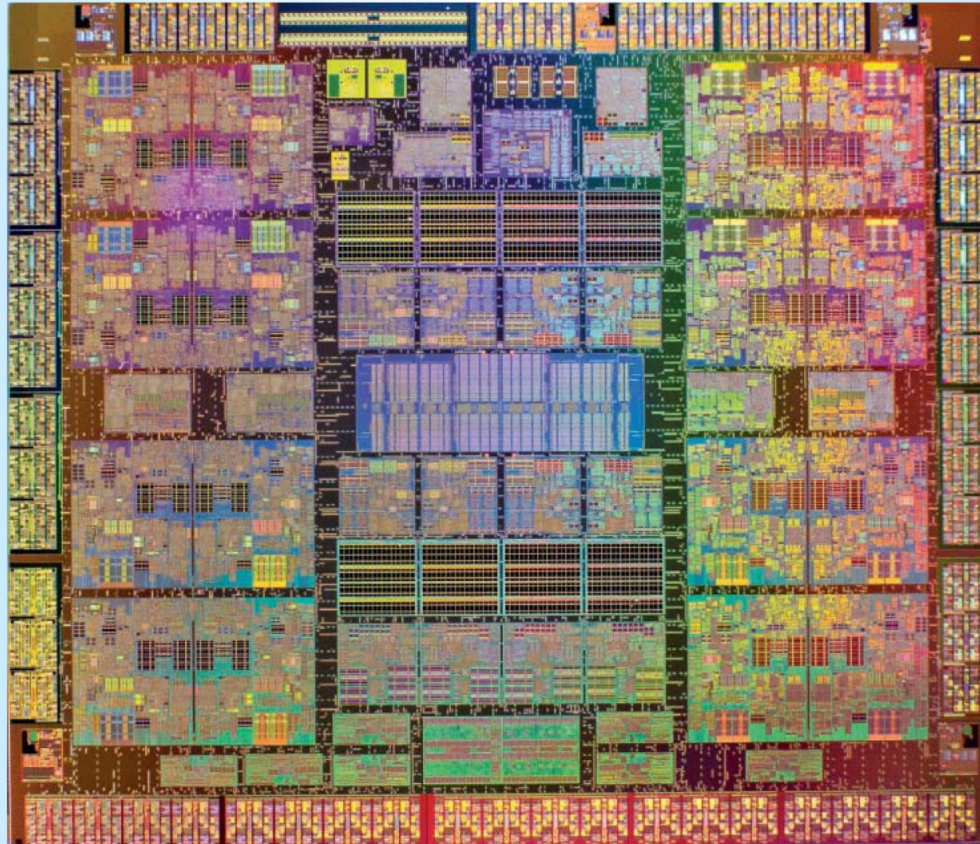
▶ UltraSparc T2 (Niagara 2)

- 8 processadores baseados no instruction set Sparc-V9
 - Cada processador é capaz de executar 8 threads concorrentemente (total de 64 threads)
 - “Server on a Chip”
 - PCI express
 - Duas portas 10 Gigabit Ethernet
 - 4 controladores dual-channel FBDIMM
- Pipeline de 8 estágios (1.6Ghz)
- Cache L2 4MB
 - 8 Bancos
 - Associativa por conjunto 16
- Aplicações
 - WebServers
 - Database



Exemplos industriais de MPSoCs (2/4)

► UltraSparc T5 (Oracle)

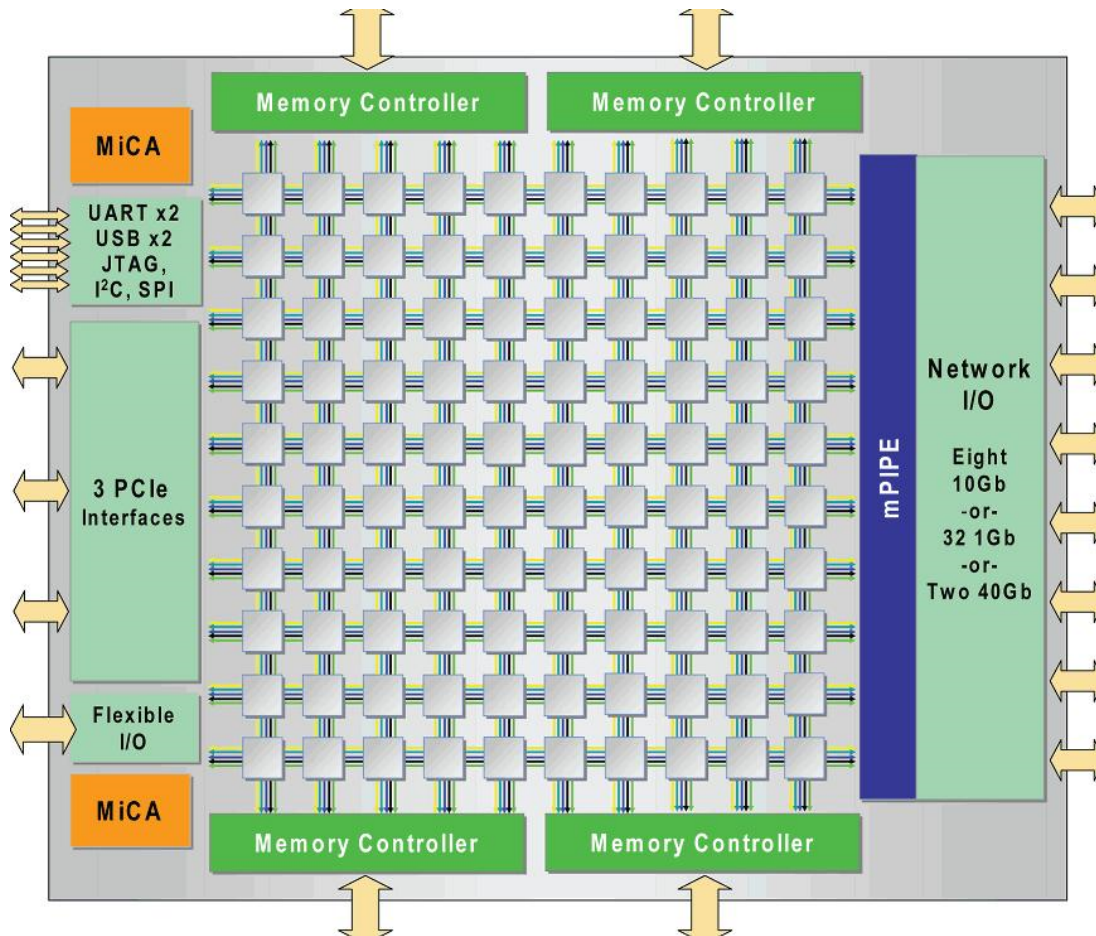


Fonte: IEEE Micro
Março/Abril 2013

Figure 1. Sparc T5 die micrograph. The crossbar is in the center with the L3 cache banks above and below it. On each side of the crossbar are the memory control units with the cores above and below them. Below the lower L3 cache banks are the coherence units, and above the upper L3 cache banks is the I/O subsystem.

Exemplos industriais de MPSoCs (3/4)

Tile GX– Tileria (100 Núcleos)

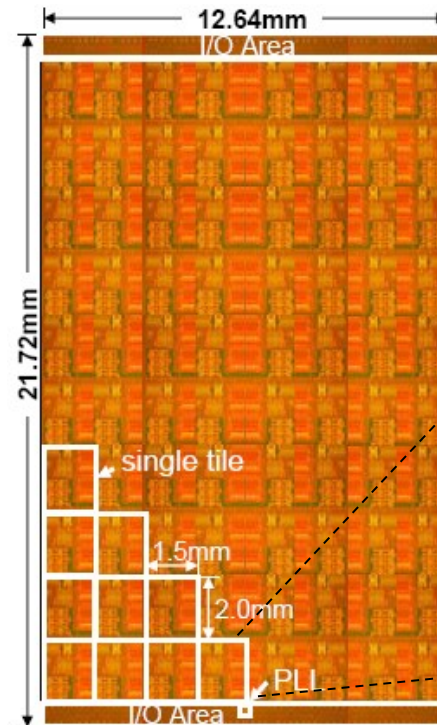


- **100 cores on a single chip**
 - 40 nm technology
- **64-bit VLIW processors**
 - Up to 3 instructions/cycle
 - 3-stage pipeline
 - Up to total of 750 BOPS
 - 32K L1i cache, 32K L1d cache, 256K L2 cache per tile
- **5 on-chip Mesh networks**
 - Over 200 Tbps
- **1 to 1.5 GHz clock frequency**
- **Supports SMP Linux and virtualization**
- **Other members: 16, 32, 64 tiles**

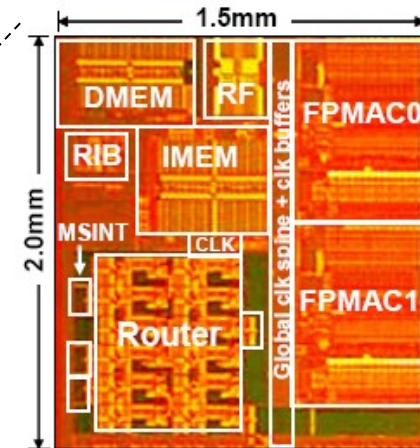
Exemplos industriais de MPSoCs (4/4)

Intel – (80 Núcleos)

- MPSoC Homogêneo
- Baseado em NoC
- 1 trilhão de operações de ponto flutuante por segundo
- Consumo de potência: 62 Watts
- Baseado em uma arquitetura VLIW
- Exploração do ILP via compilador



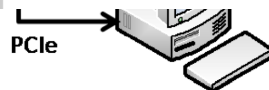
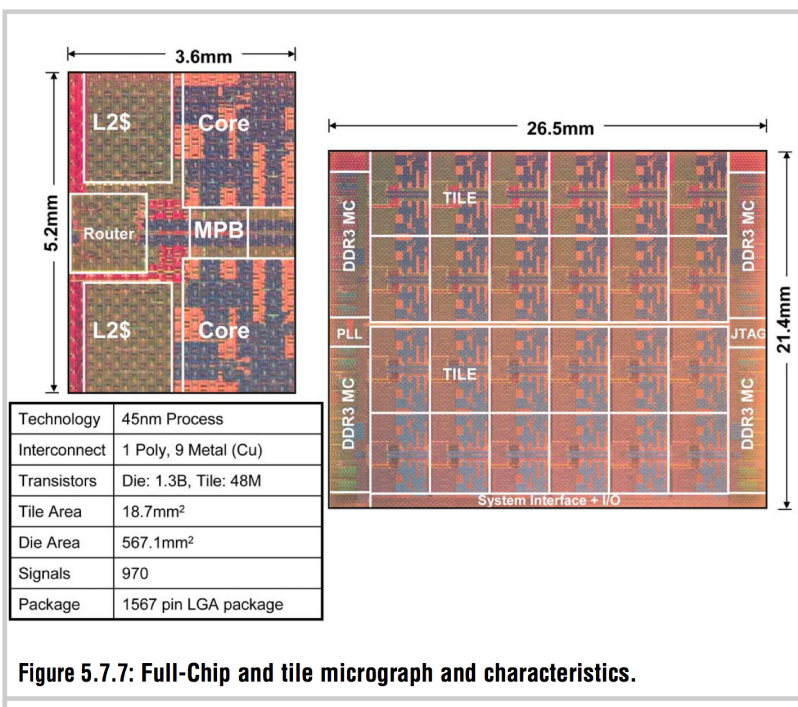
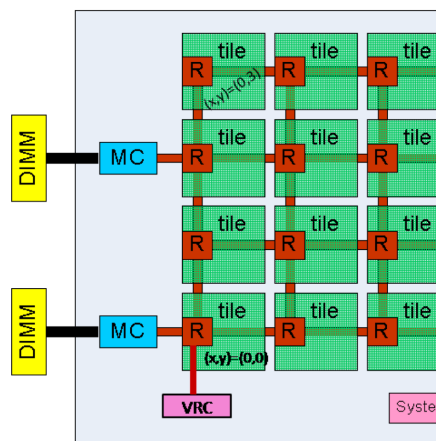
MPSoC proposto pela Intel (2007)



Simplicidade dos Cores:
Remoção do HW superescalar

Intel Single-chip Cloud Computer (SCC)

- 24 “tiles” with two IA cores per tile
- A 24-router mesh network with 256 GB/s bisection bandwidth
- 4 integrated DDR3 memory controllers
- Hardware support for me



System Architecture

■ Processor

- Dual-core ARM® Cortex™-A9 MPCore™ processor
- 4,000 MIPS (up to 800 MHz per core)
- NEON coprocessor with double-precision FPU
- 32-KB/32-KB L1 caches per core
- 512-KB shared L2 cache

■ Multiport SDRAM controller

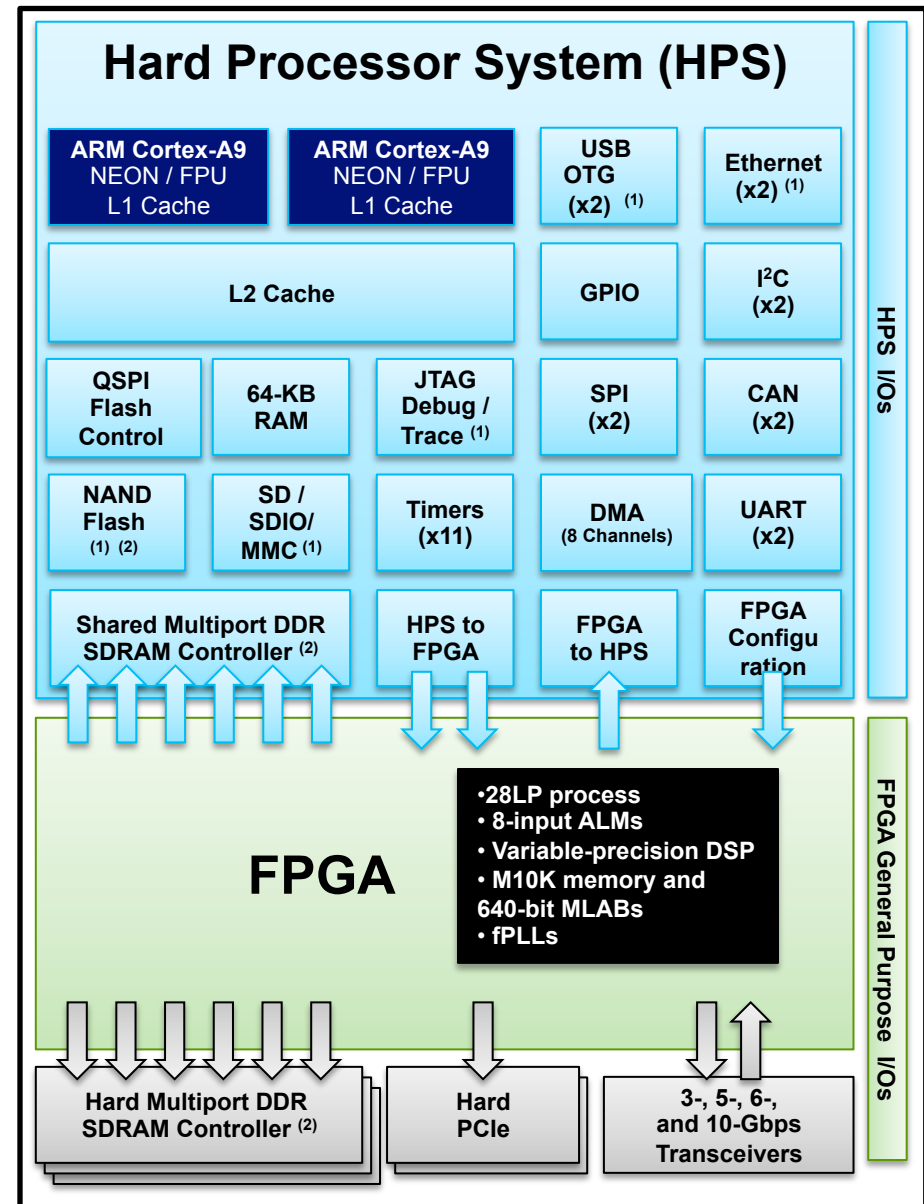
- Up to 533-MHz DDR3 and LPDDR2
- Up to 400-MHz DDR2
- Up to 200-MHz Mobile DDR
- Integrated ECC support

■ High-bandwidth on-chip interfaces

- > 125-Gbps HPS-to-FPGA interface
- > 125-Gbps FPGA-to-SDRAM interface

■ Cost- and power-optimized FPGA fabric

- Lowest power transceivers
- Up to 1,600 GMACS, 300 GFLOPS
- Up to 25Mb on-chip RAM
- More hard intellectual property (IP): PCIe® and memory controllers



Notes:

(1) Integrated direct memory access (DMA)

(2) Integrated ECC



HeMPS (PUCRS)

- **MPSoC Homogêneo**

- Componentes:

- Noc Hermes
 - Plasma-IP

- Mestre e Escravo:

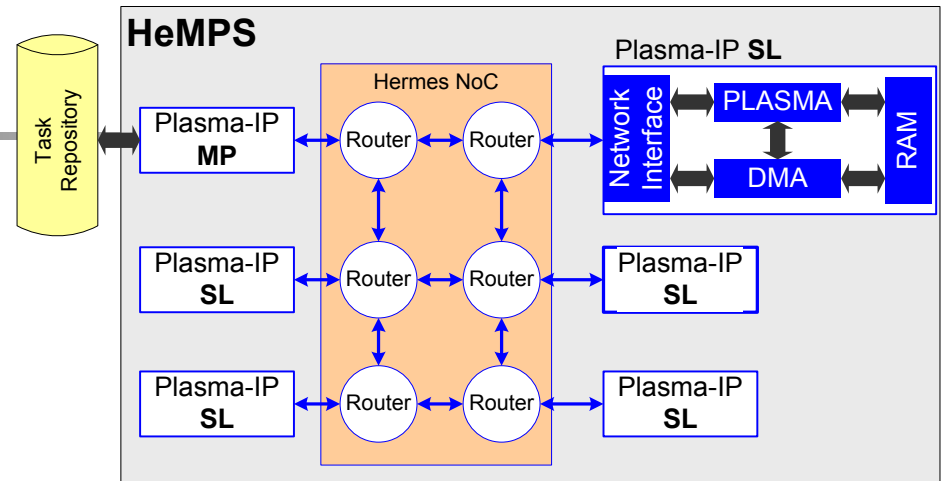
- Plasma IP-MP: gerência dos recursos do sistema.
 - Plasma IP-SL: executam as tarefas das aplicações

- Repositório de tarefas

- Externo ao MPSoC
 - Permite mapeamento dinâmico de tarefas
 - Acessado somente pelo Plasma IP-MP

- Cada processador executa um microkernel

- Suporte a mapeamento dinâmico de tarefas
 - Suporte a execução multitarefa → escalonamento round robin



HeMPS

■ PLASMA-IP

■ Processador

- Plasma (opencores.org)
- 32 bits MIPS R3000-like CPU
- No MMU / No cache
- Suporta interrupções e *timers*

■ RAM

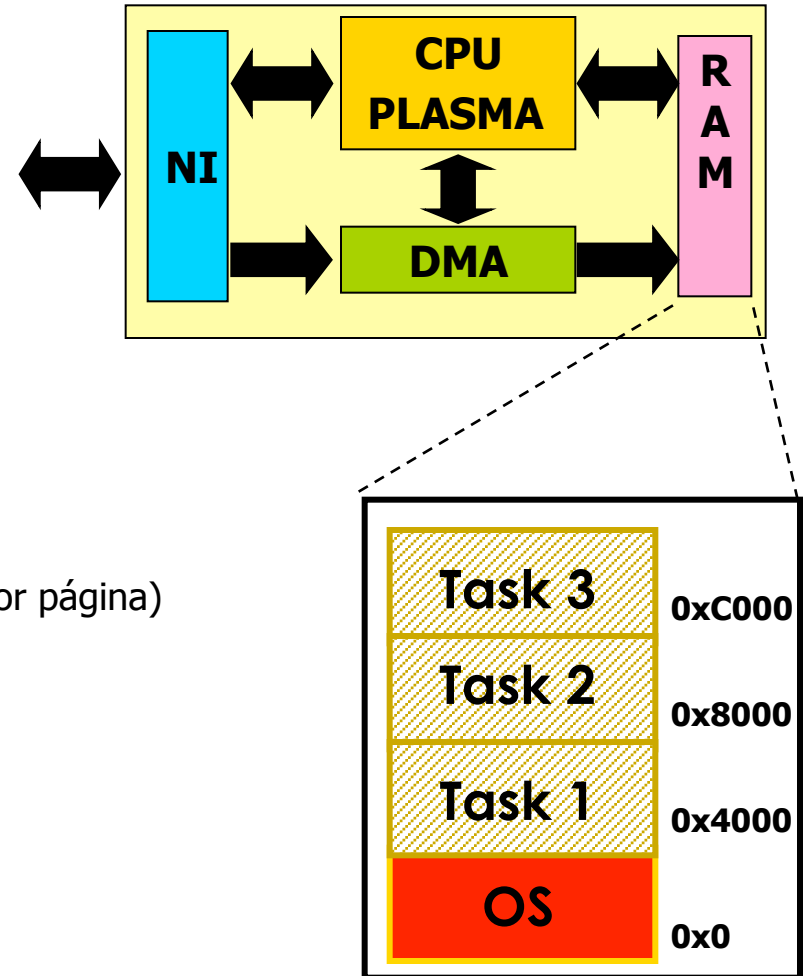
- 64 KB
- Organizada em 4 pages (4 K 32-bit palavras por página)

■ DMA

- Separa comunicação de comunicação
- Responsável por receber tarefas, e realizar a comunicação com a NoC

■ NI – Network Interface

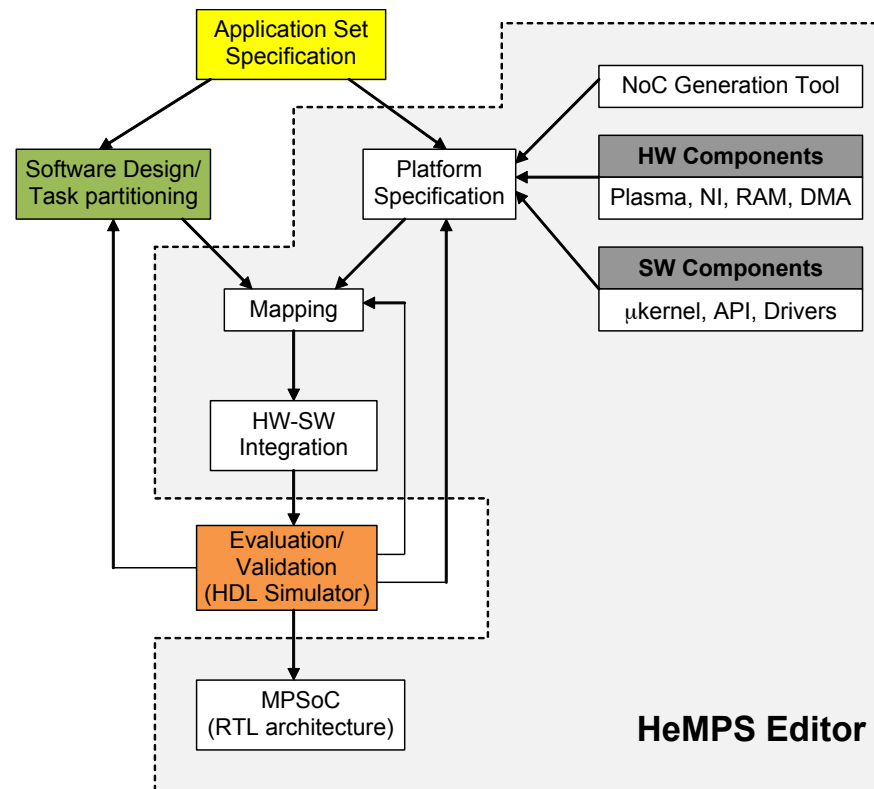
- Interface com a rede



HeMPS Framework Editor

■ Design Flow

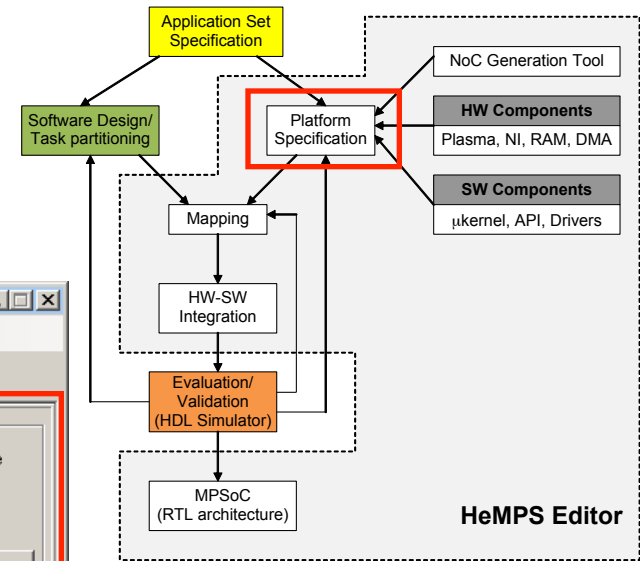
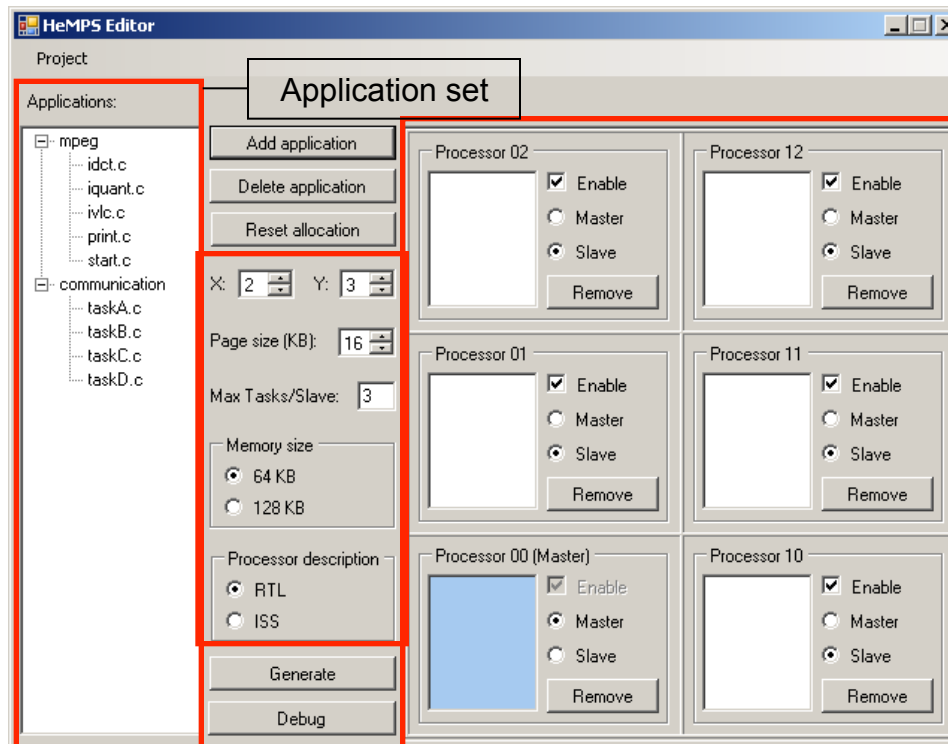
- Platform-based design
- Several steps covered by the HeMPS Editor



HeMPS Framework Editor

■ HeMPS Editor

- Quick platform customization
- Automates the platform generation
- Provides abstract models (Plasma/RAM)



HeMPS configuration

HeMPS Editor

HeMPS Framework Editor

■ HeMPS Editor

- Quick platform customization
- Automates the platform generation
- Provides abstract models (Plasma)

The screenshot displays the HeMPS Editor interface. On the left, the 'Project' pane shows a tree structure of applications: 'mpeg' (containing 'iquant.c', 'ivlc.c', and 'communication'), 'taskC.c', and 'taskD.c'. The 'communication' application is selected. The 'Add Application' button is highlighted. The 'Reset Allocation' button is also visible. The 'Page Size (KB)' is set to 16. The 'Max Tasks/Slave' is set to 2. The 'Memory Size' is set to 64 KB. The 'Processor Description' is set to ISS. The 'Generate' and 'Debug' buttons are highlighted with red boxes. The 'Report' window on the right shows the execution log for various processors. The 'Application Set Specification' and 'NoC Generation Tool' labels are present at the top right. The 'Report' window contains the following text:

Processor 02
taskA.c (6) | taskB.c (7) | iquant.c (5)
Communication task A started.
4291
8094
Communication task A finished.

Processor 12
start.c (8)
MPEG Task A start: Send block
6368
T1
8258
T2
10224

Processor 01
ivlc.c (1) | print.c (2)
MPEG Task B start: IVLC
11858
359007
End Task B - MPEG

Processor 11
taskD.c (4)
Communication task D started.
16110
10
11
12
13

Processor 00 (Master)
Kernel
Master strikes again.

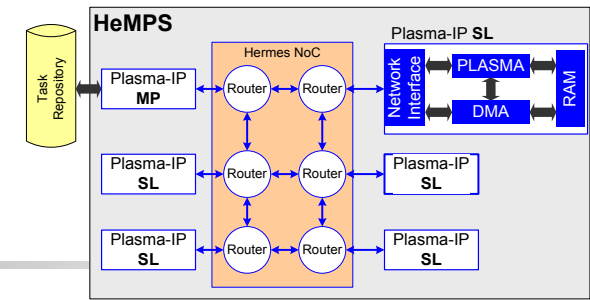
Processor 10
taskC.c (3) | idct.c (0)
Communication task C started.
9184
26338
Communication task C finished.

Processor 00 (Master)
idct.c
print.c
taskA.c
taskB.c

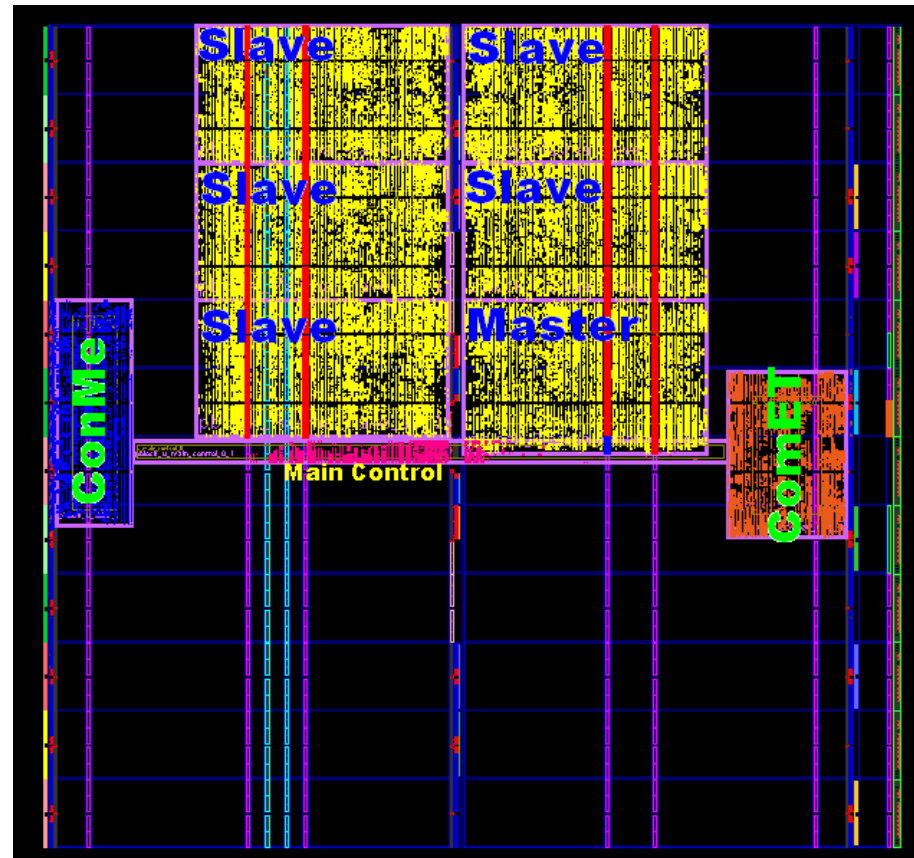
or

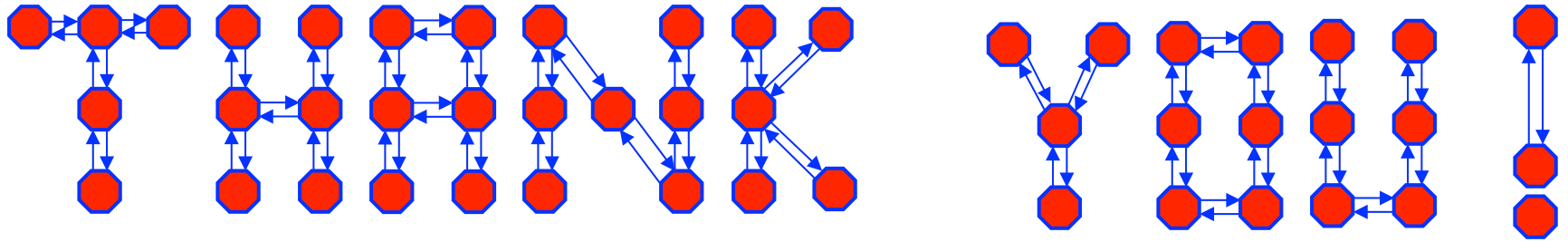
HeMPS

- Operacional em FPGA
 - Interface com o PC: ethernet
 - Repositório em memória DDR



HeMPS 2x3 em FPGA
Virtex 5 LX 300:





Questions?

fernando.moraes@pucrs.br