# ALGORÍTMOS PARALELOS
## (Aula 2)

**Neyval C. Reis Jr.**

**OUTUBRO/2004**


**Laboratório de Computação de Alto Desempenho**

**DI/UFES**

UFES
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

# Programa do Curso

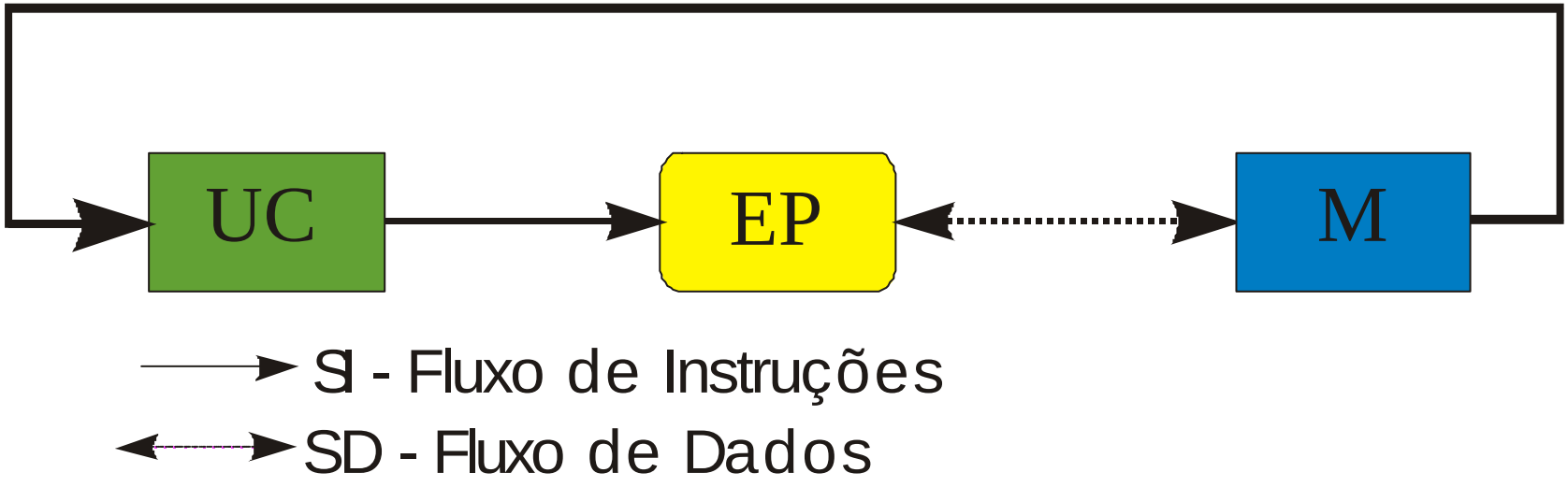

1. Introdução
2. Arquitetura de Computadores
3. Arquiteturas de Sistemas Paralelos
4. Computação de Alto Desempenho
5. Programação Paralela (modelos e paradigmas)
6. Análise de Desempenho e Instrumentação
7. Aplicações

UFES
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

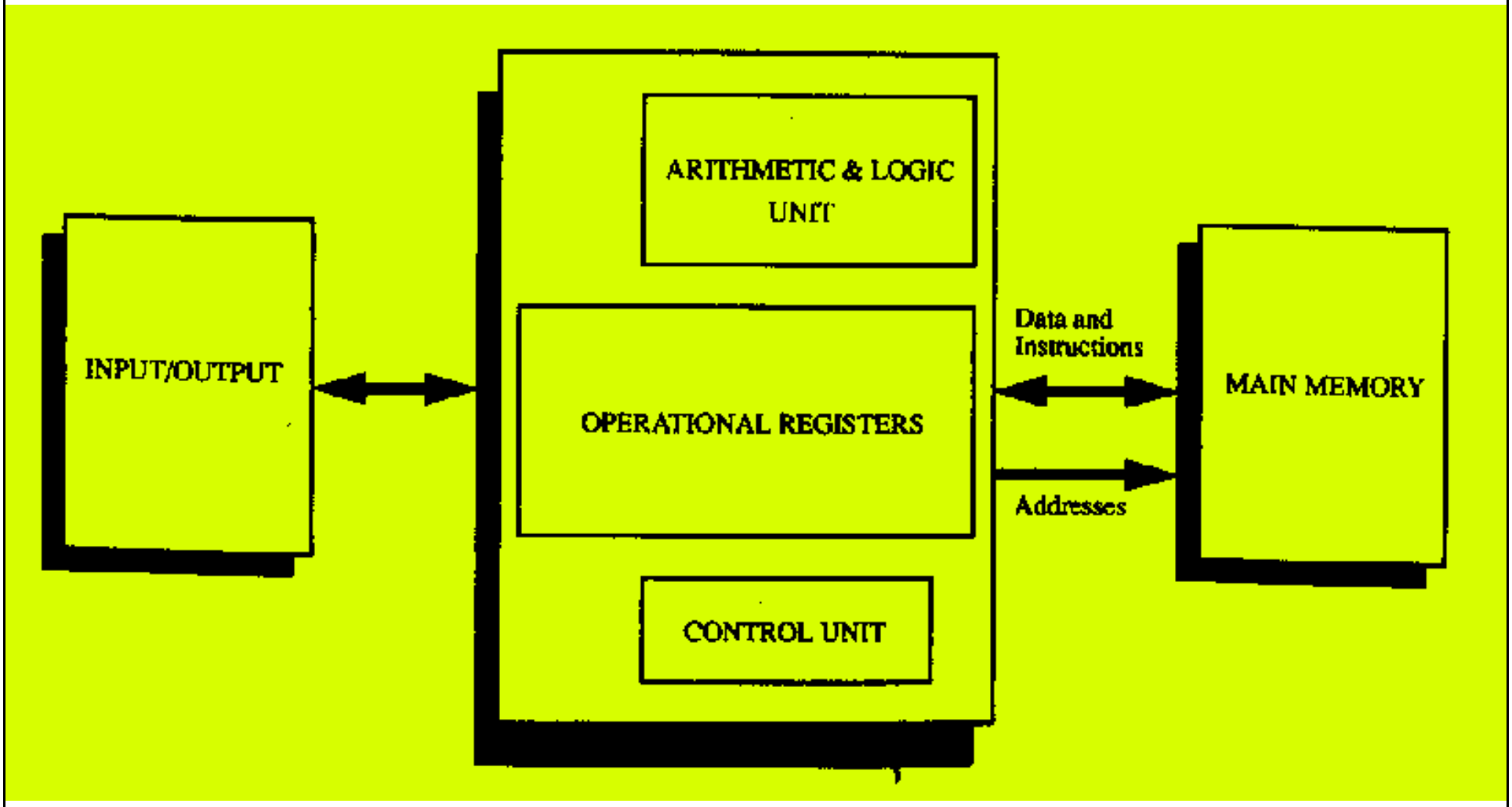
# Arquitetura de processadores e Processamento de Alto Desempenho

# Modelo de computador proposto por Von-Neumann

UC – Unidade de controle
EP – Elemento de processamento
M - Memória

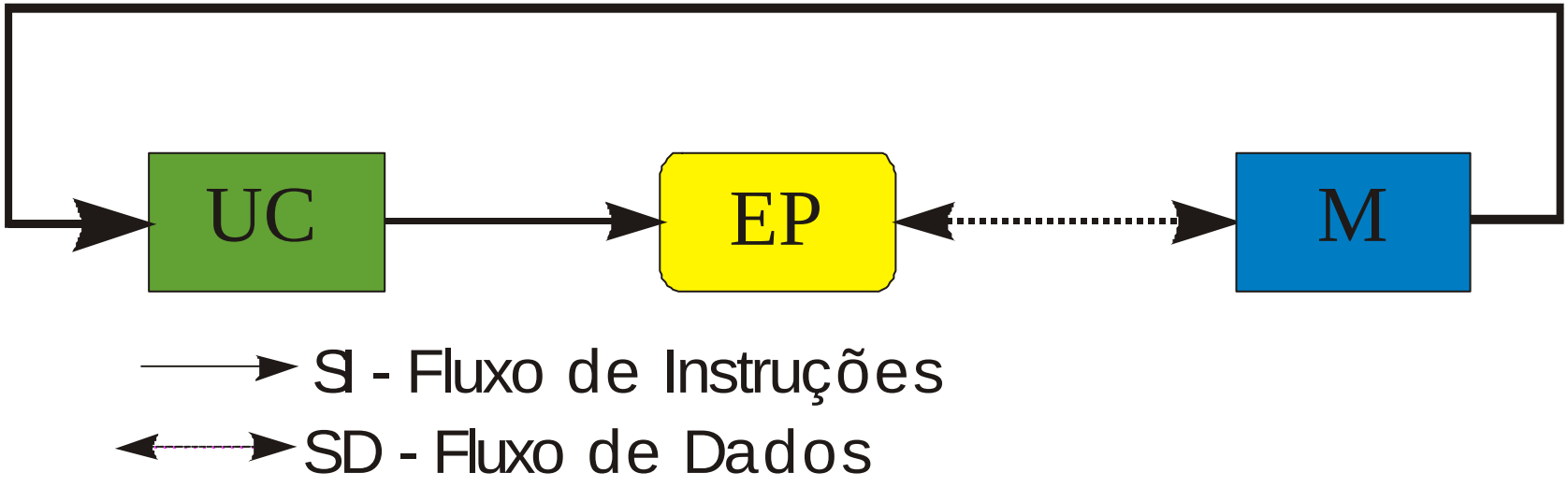


# Arquitetura de um microprocessador



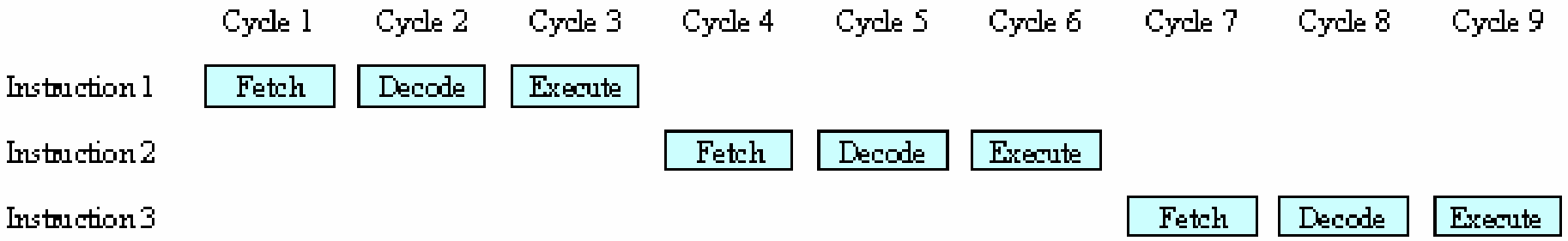
# Modelo de computador proposto por Von-Neumann

UC – Unidade de controle
EP – Elemento de processamento
M - Memória

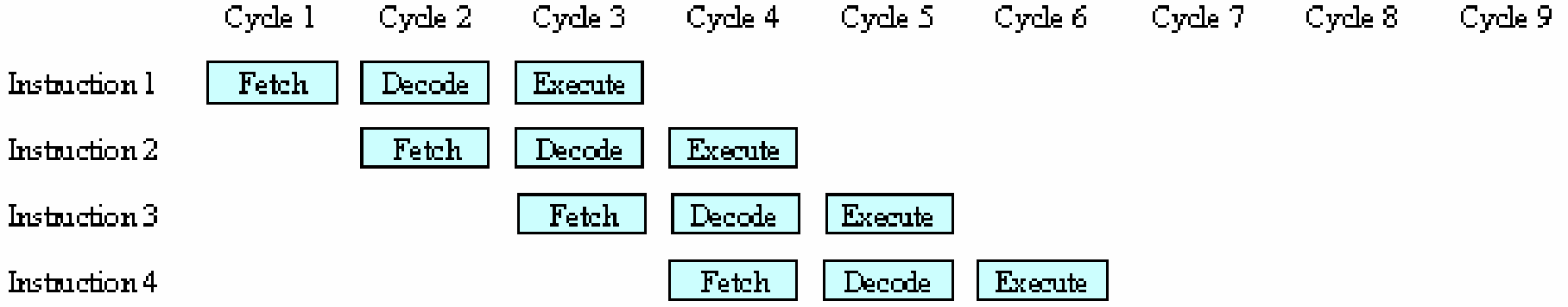


# Processadores Escalares

## Non-pipelined Processors

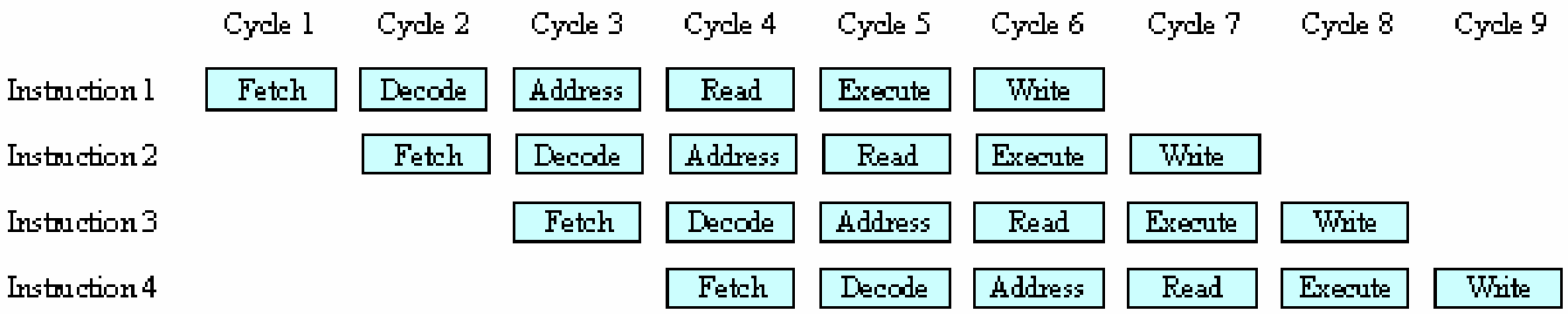



## Three-stage Pipelined Processors



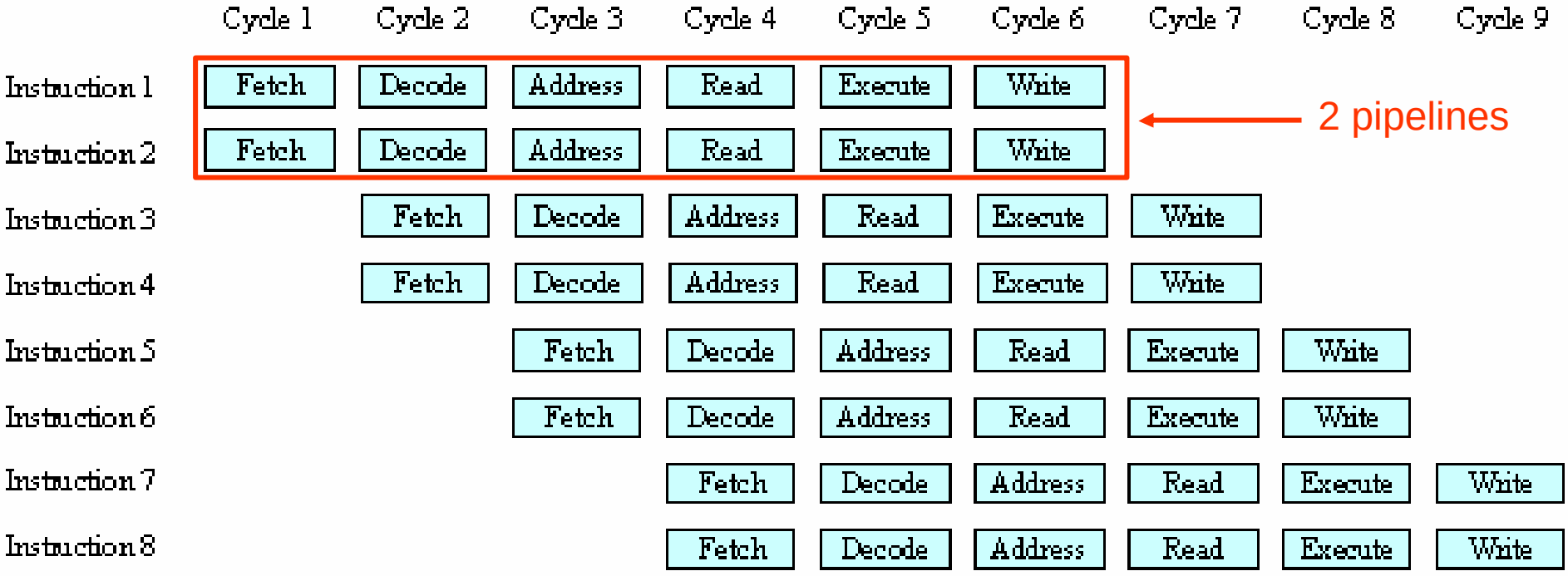
# Processadores Escalares

## Extended-pipelined Processors



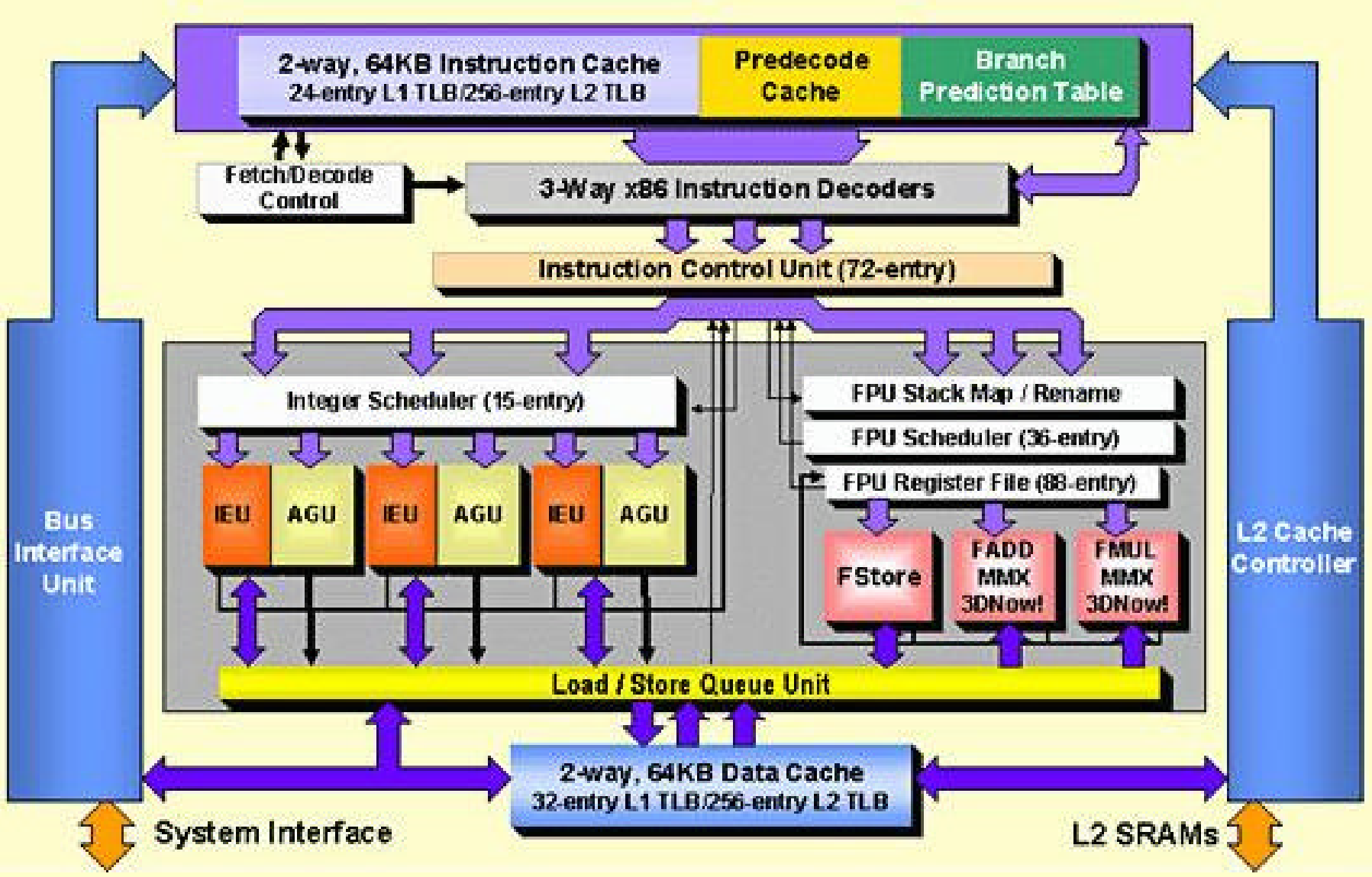
# Processadores Superescalares

## Six-stage Level 2 Superscalar Processors

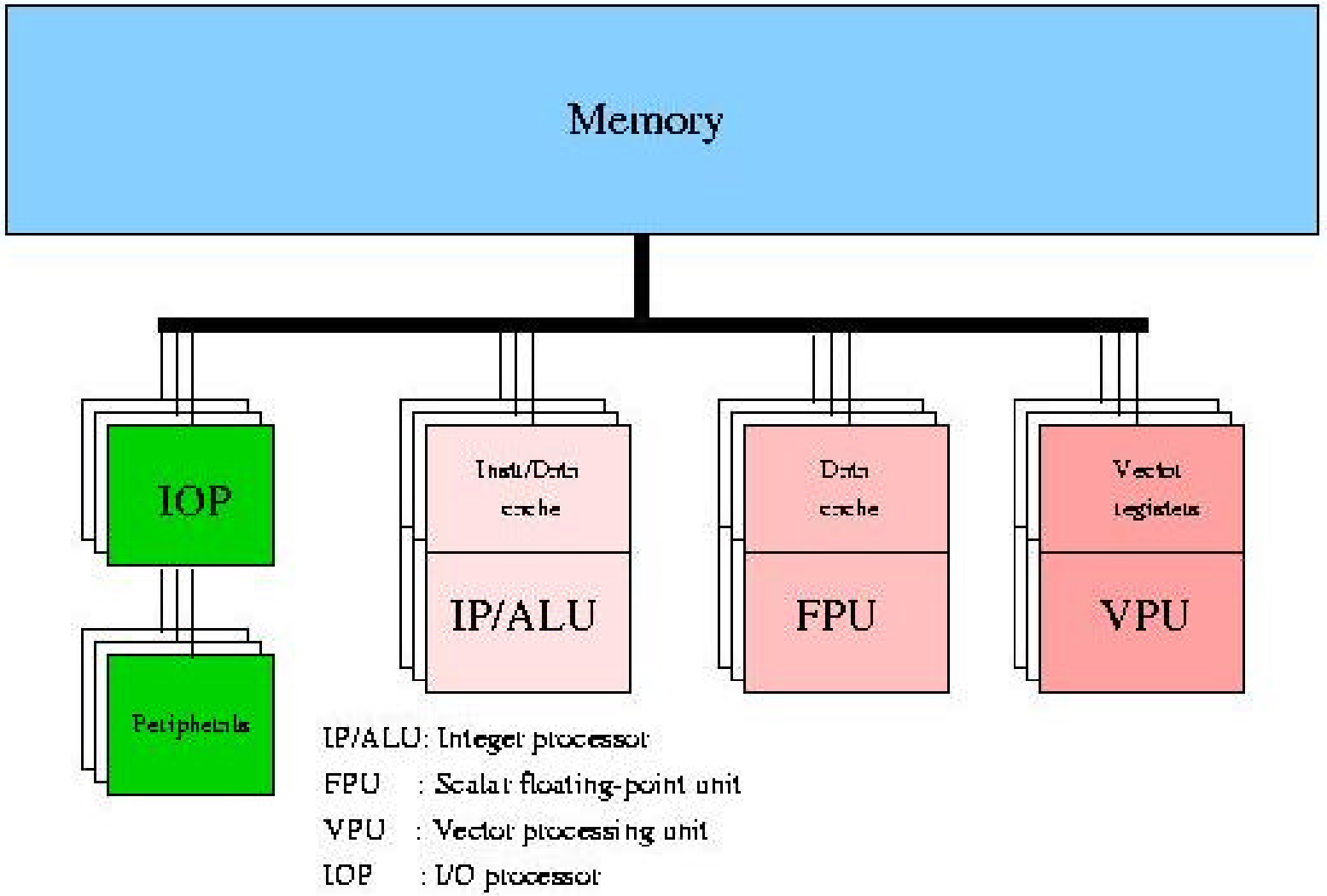



Dependências de dados e desvios podem reduzir consideravelmente
o desempenho de arquiteturas superescalares

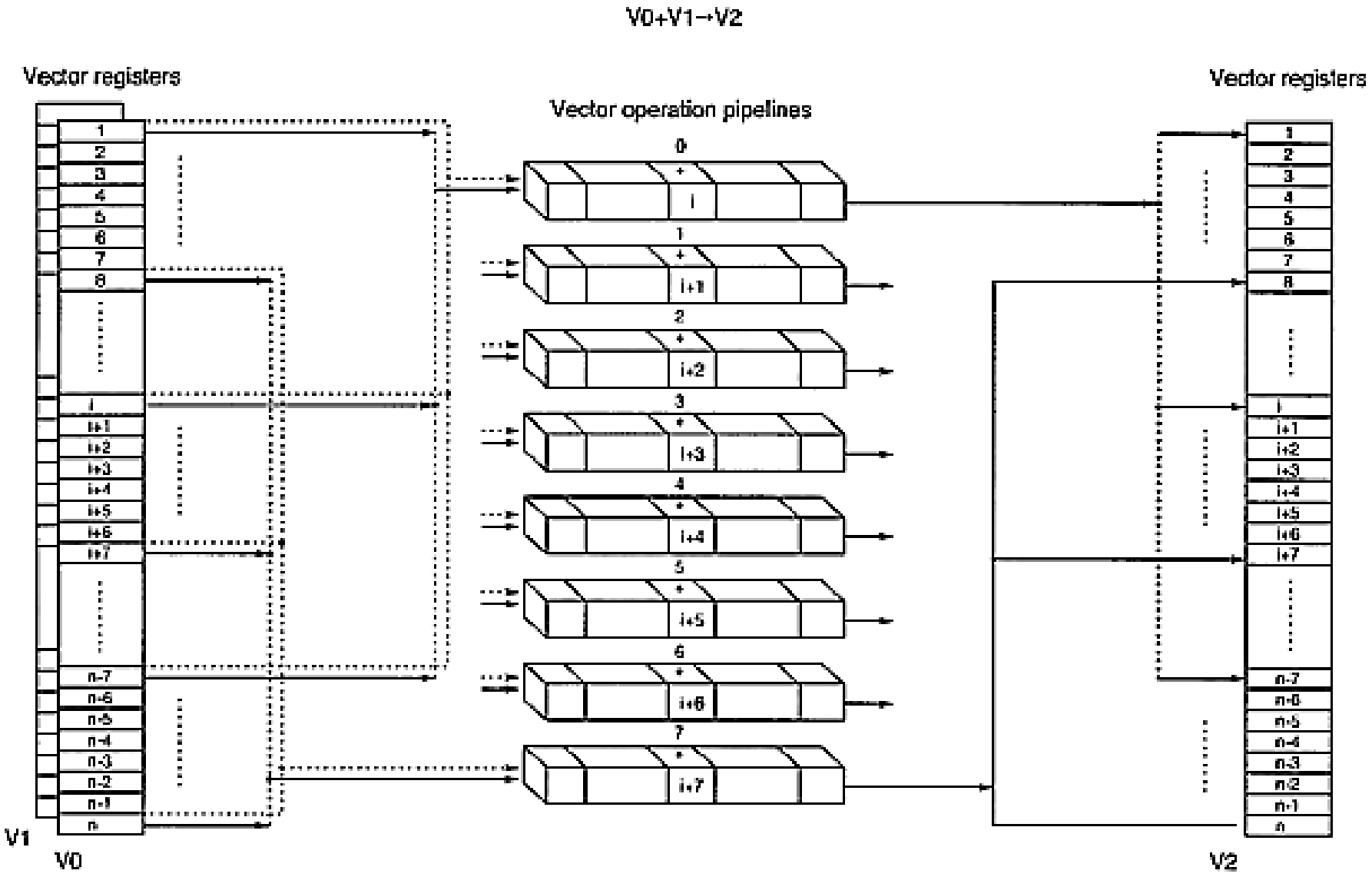
# Processadores Superescalares (exemplo: AMD ATLHON)



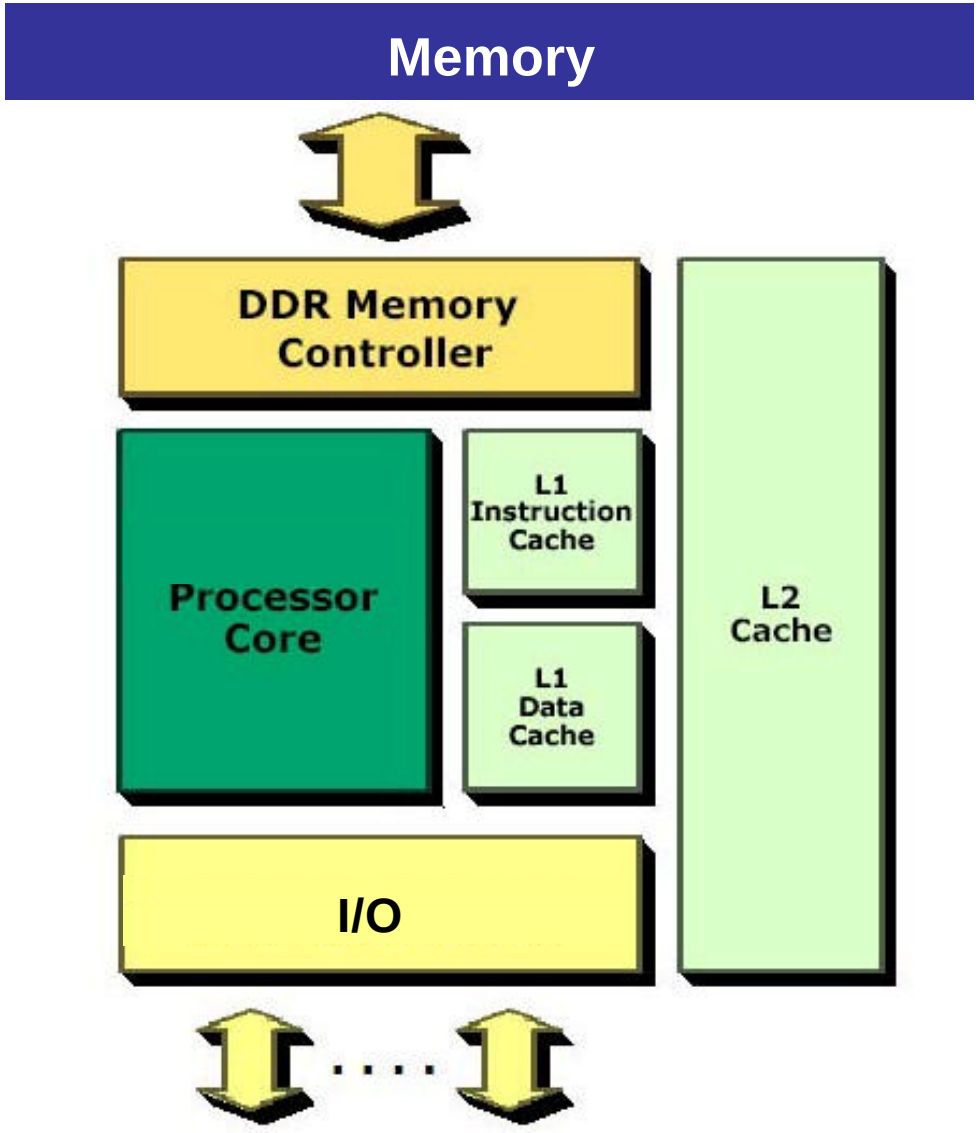
# Processadores Vetoriais



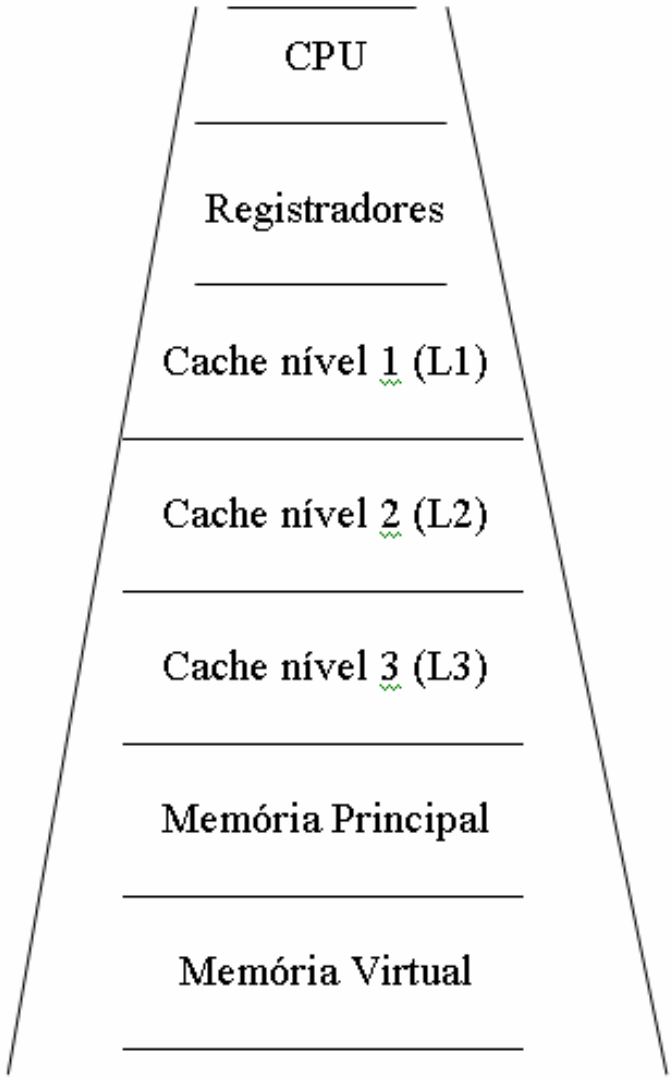
# Processadores Vetoriais



# Hierarquia de Memória



# Hierarquia de Memória

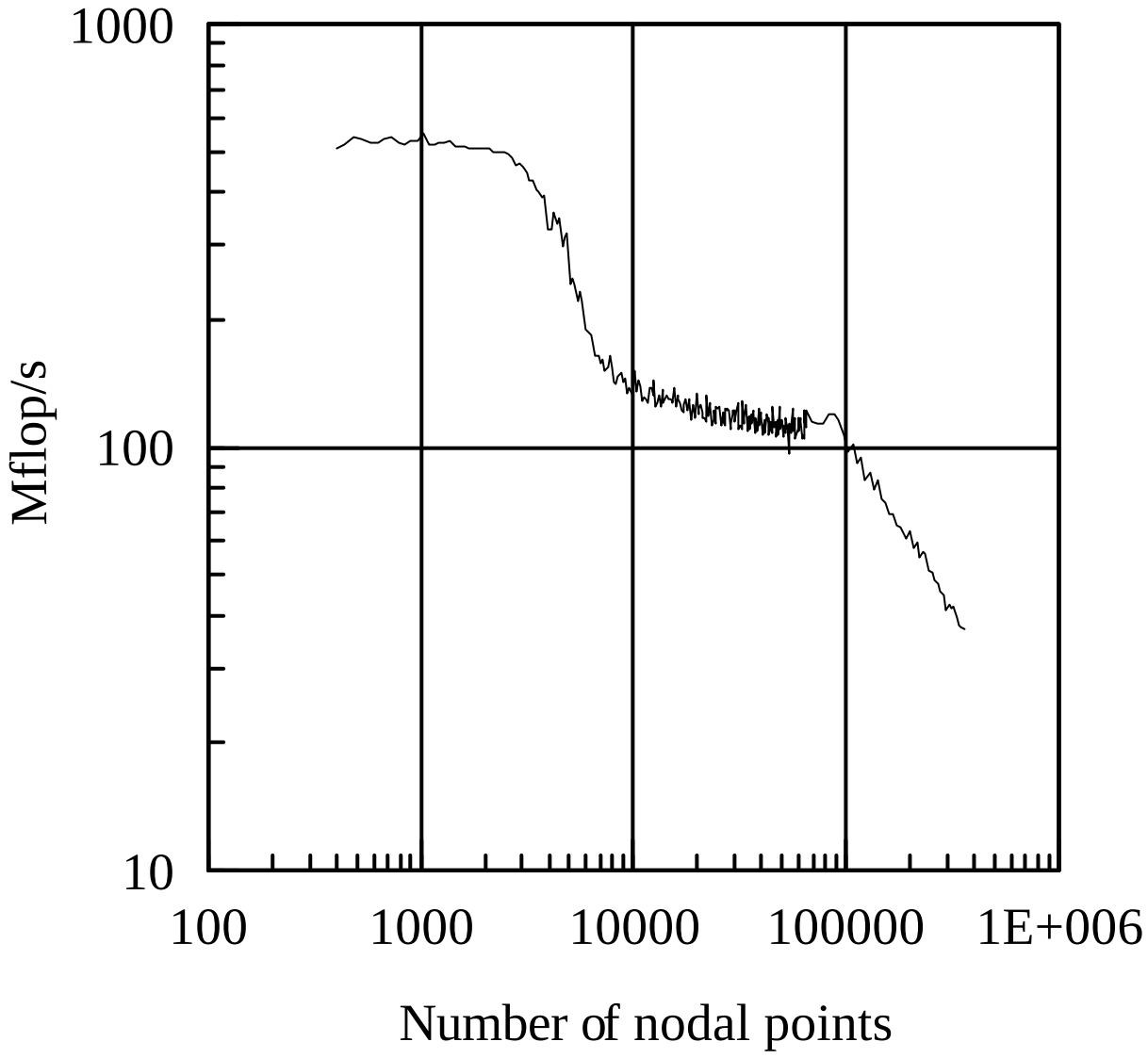



| | Capacidade | Velocidade de transferência de dados | Latência |
|---|---|---|---|
| ← | 256 Bytes | 24000 MB/s | 2 ns |
| ← | 8 Kbytes | 16000 MB/s | 2 ns |
| ← | 96 Kbytes | 8000 MB/s | 6 ns |
| ← | 2 Mbytes | 888 MB/s | 24 ns |
| ← | 1.5 Gbytes | 1000 MB/s | 112 ns |

Tabela 1 – Performance em Mflops obtida na solução de um sistema linear de equações em função do tamanho das matrizes de dados. Resultados obtidos por Weiβ et al. (2000), utilizando um processador Alpha Digital 21164 com performance máxima de 1000 Mflops.

| Tamanho da matriz | 32 x 32 | 64 x 64 | 128 x 128 | 256 x 256 | 512 x 512 | 1024 x 1024 | 2048 x 2048 |
|---|---|---|---|---|---|---|---|
| Espaço de memória ocupado pela estruturas de dados | 17 KB | 66 KB | 260 KB | 1 MB | 4 MB | 16 MB | 64 MB |
| Mflops | 365.1 | 444.6 | 188.8 | 180.0 | 66.4 | 58.5 | 56.2 |

**MFLOPS obtained for a single processor simulation as a function of the number of nodal points.**



## Otimização

1. Evitar desvios de execução
2. Uso eficiente da memória hierarquia de memória
3. Evitar divisões
4. Permitir “Loop-unrolling”

# Programa do Curso


1. Introdução
2. Arquitetura de Computadores
3. Arquiteturas de Sistemas Paralelos
4. Computação de Alto Desempenho
5. Programação Paralela (modelos e paradigmas)
6. Análise de Desempenho e Instrumentação
7. Aplicações

UFES
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO

# Arquiteturas de

# Sistemas Paralelos

# Classificação de arquiteturas de processamento paralelo

## 2 formas principais de classificação:

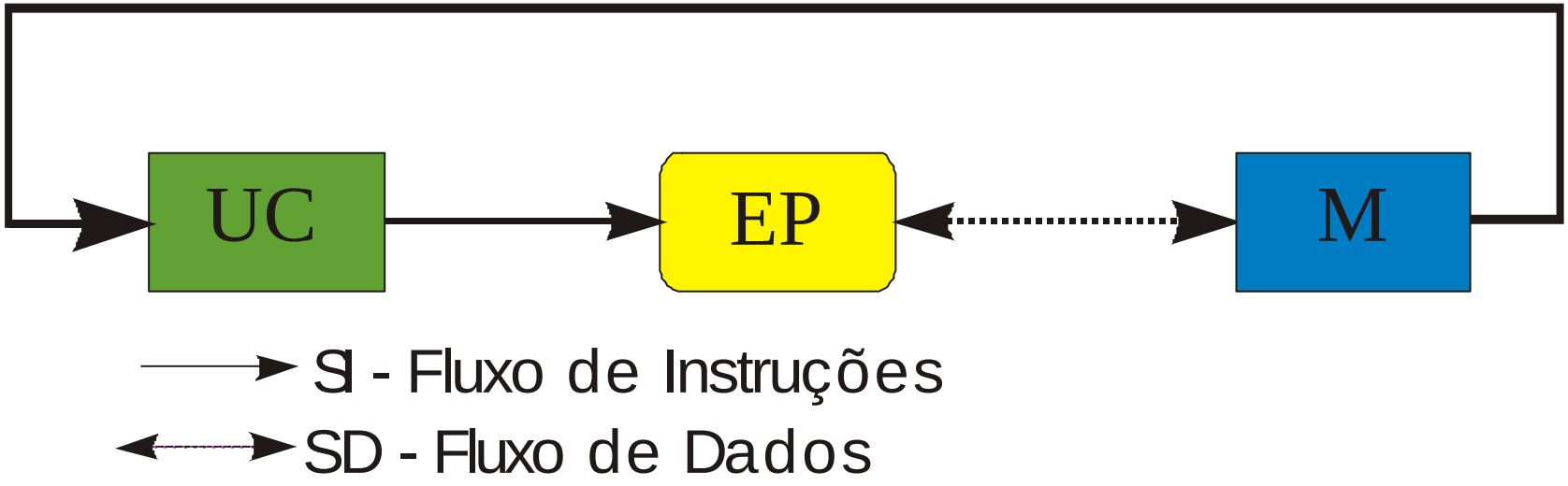
1. **Fluxo de dados e instruções (Flynn)**
2. **Compartilhamento de memória**

# Classificação de Flynn segundo fluxos de instruções e dados

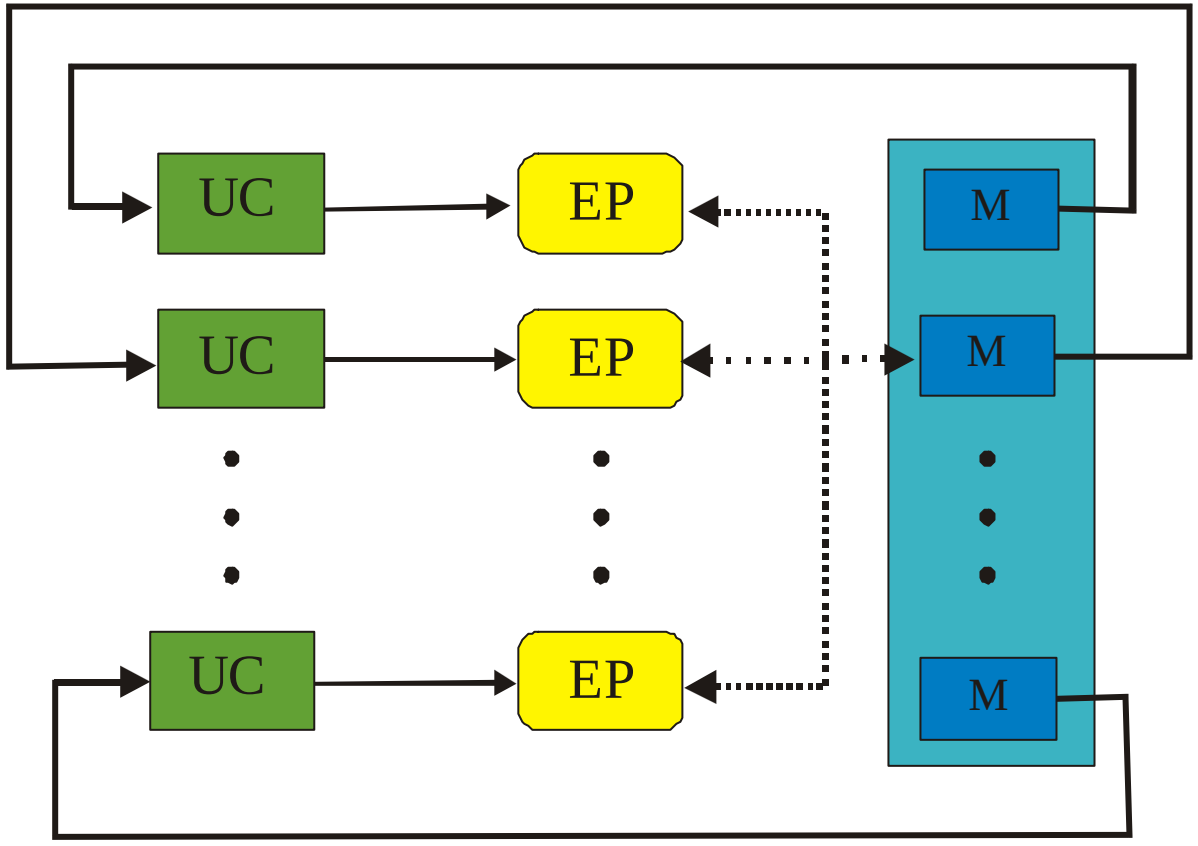
| | **SD** (Single Data) | **MD** (Multiple Data) |
|---|---|---|
| **SI** (Single Instrution) | **SISD**<br>Máquinas von Neumann Convencionais, PC's e estações de trabalho | **SIMD**<br>Máquinas Vetoriais (ICL Distributed Array Processor – DAP, Thinking Machine Corporations – CM-200) |
| **MI** (Multiple Instrution) | **MISD**<br>Sem representante | **MIMD**<br>Multiprocessadores e Multicomputadores (nCUBE, Intel Paragon, Cray T3D) |

# Diagrama da classe SISD

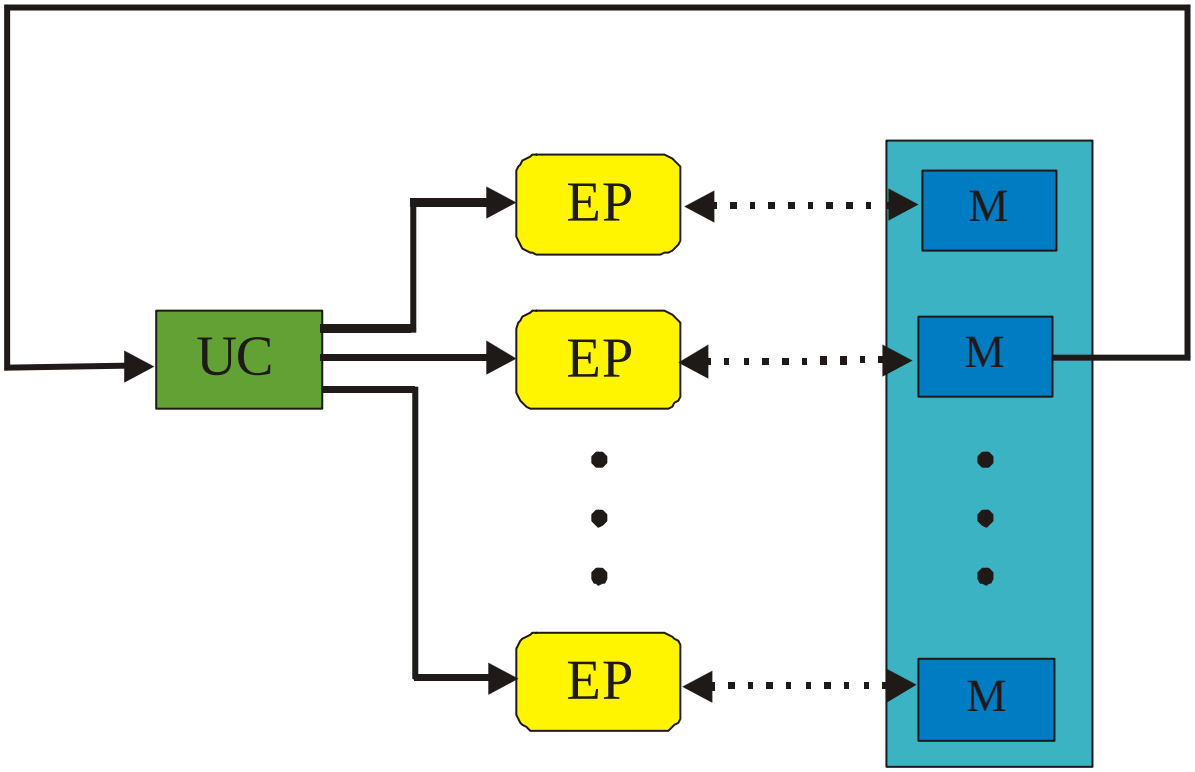
UC – Unidade de controle
EP – Elemento de processamento
M - Memória



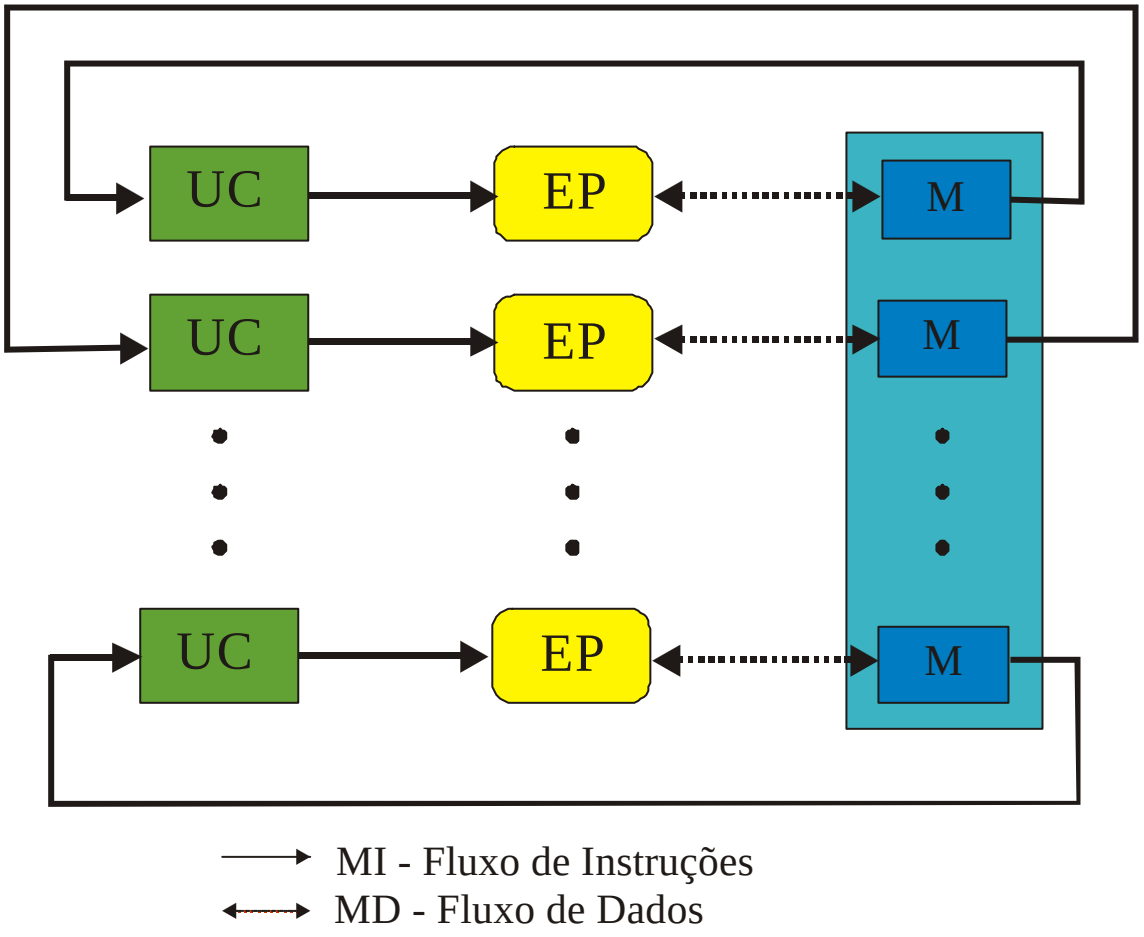
# Diagrama da classe MISD



# Diagrama da classe SIMD



# Diagrama da classe MIMD

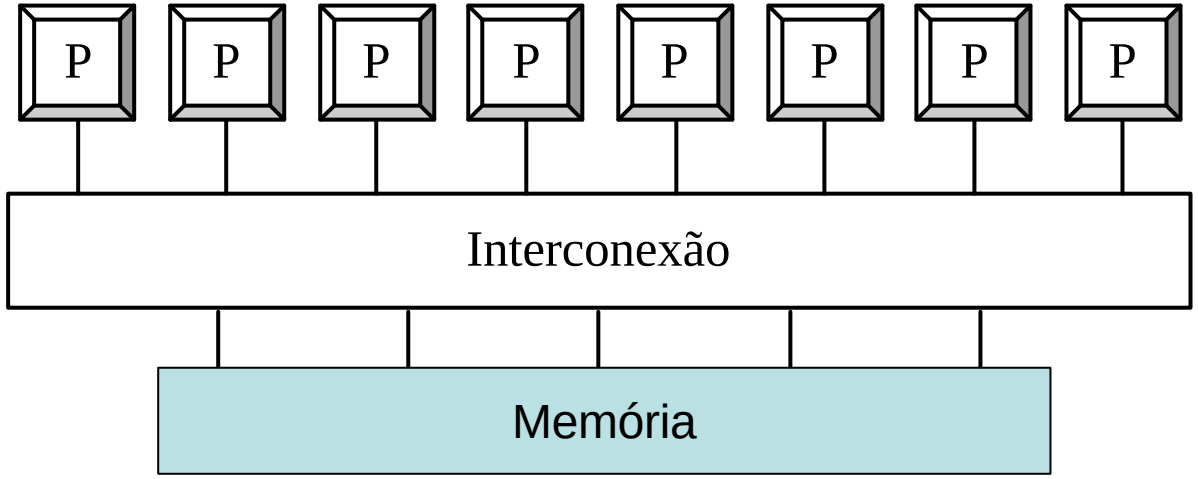


## Classificação segundo o compartilhamento de memória

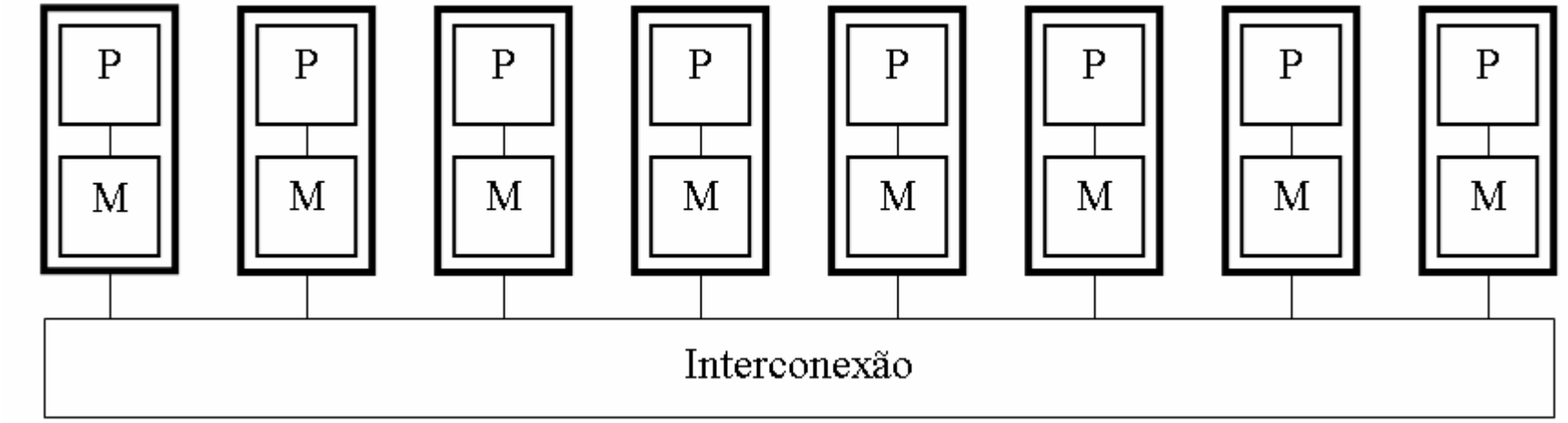
Esse é um outro critério que classifica máquinas paralelas segundo sua memória.

- Nas máquinas com **memória compartilhada** (*shared memory*) existe apenas um espaço de endereçamento, que é usado de forma implícita para a comunicação entre os processadores, com operações de *load* e *store*.
- O outro caso é quando existem **múltiplos espaços de endereçamento privados** (*multiple private address spaces*), um para cada processador (**memória distribuída**). A comunicação tem que ser explícita, através do uso de troca de mensagens com operações *send* e *receive*.

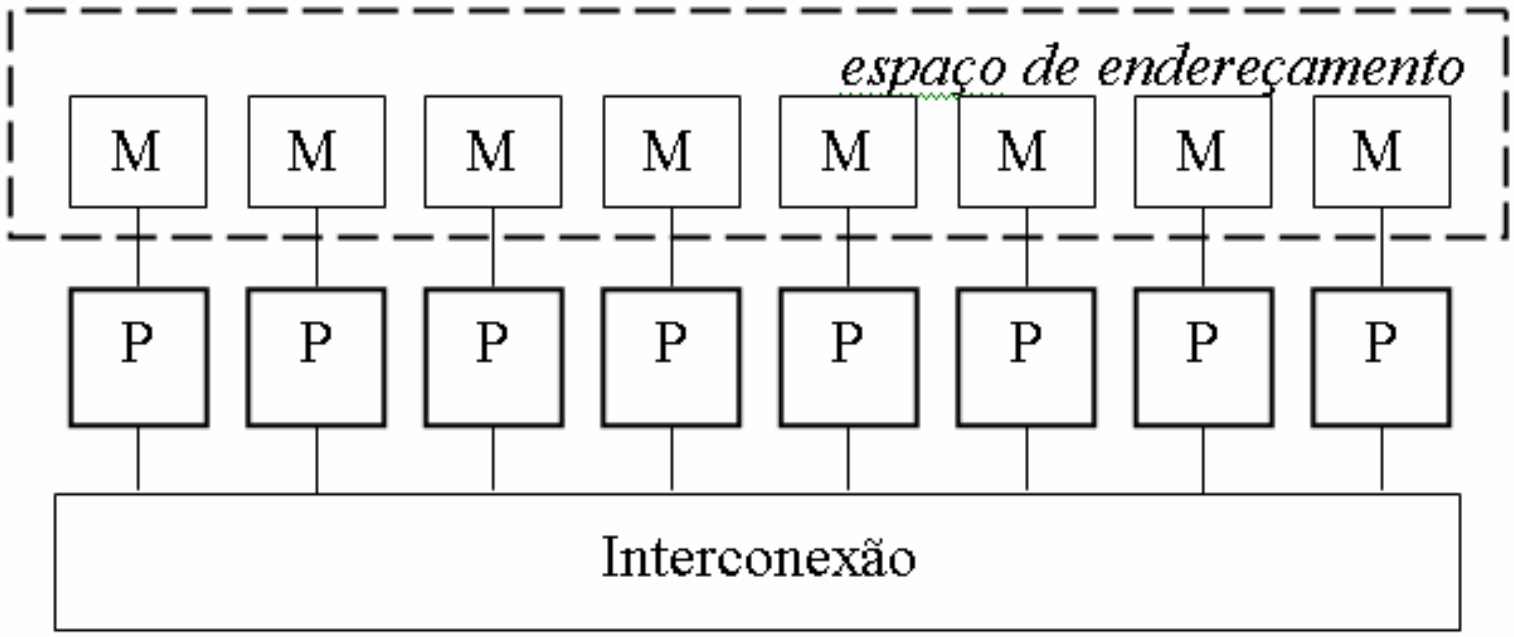
# Memória Compartilhada



# Memória Distribuída



# Memória Compartilhada Distribuída (DSM)




- Endereçamento Único (memória compartilhada)
- Fisicamente distribuída
- Tempo de acesso não uniforme (NUMA ✍ Non-Uniform Memory Access)
- Implementação através de Hardware ou Software

# COMPARAÇÃO ENTRE OS MODELOS FÍSICOS PARA CONSTRUÇÃO DE MÁQUINAS PARALELAS

| Tipo | PVP | SMP | MPP | DSM | NOW |
|---|---|---|---|---|---|
| Estrutura | SIMD | MIMD | MIMD | MIMD | MIMD |
| Comunicação | memória compartilhada | memória compartilhada | troca de mensagem | memória compartilhada | troca de mensagem |
| Interconexão | Crossbar | barramento ou crossbar | rede específica | rede específica | rede comum |
| Número de nós | ~10 | ~50 | Acima de 100 | 10 a 2000 | 4 a 10000 |
| Endereçamento | único | único | múltiplo | único | múltiplo |
| Processador | Específico | comum | comum | específico ou comum | comum |
| Acesso à memória | UMA | UMA | NORMA | NUMA | NORMA |

NORMA – Non-remote memory access
UMA – Uniform memory access
NUMA – Non-uniform memory Access

# Processadores Vetoriais Paralelos (PVP)

Sistemas compostos de poucos processadores poderosos. A interconexão é feita, em geral, por uma matriz de chaveamento (crossbar) de alta vazão. A memória é compartilhada, e os sistemas podem ser classificados como multiprocessadores UMA. Normalmente não utilizam memória cache, usando para essa função um grande número de registradores vetoriais e um buffer de instrução. Exemplos: Cray C-90 (máximo de 16 processadores), Cray T-90 (máximo de 32 processadores), Fujitsu VPP 700 (máximo de 256 processadores). O NEC Sx-6 também é um PVP, e o Earth Simulator, que é um NEC Sx-6, é o número 1 na lista das 500 máquinas mais poderosas do mundo, possuindo 5120 processadores.

# Multiprocessadores simétricos (SMP)

Os Symmetric Multiprocessors são sistemas constituídos de processadores comerciais conectados a uma memória compartilhada, podendo também ser classificados como multiprocessadores UMA. Utilizam-se amplamente de memória cache e todos os processadores têm igual acesso ao barramento e à memória compartilhada. São mais fáceis de programar que máquinas que se comunicam pro troca de mensagens, já que a forma de programação se aproxima daquela feita em sistemas convencionais, mas tem como desvantagem o uso de um barramento de interconexão (permitindo apenas uma transação por vez). Esta limitação pode reduzir a escalabilidade desta classe de sistemas, fazendo com que sistemas comerciais estejam, geralmente, limitados a 64 processadores. Exemplos: IBM R50 (máximo de 8 processadores), SGI Power Challenge (máximo de 36 processadores), SUN Ultra Enterprise 10000 (máximo de 64 processadores) e HP/Convex Exemplar X-Class (máximo de 32 nós de 16 processadores cada).

## Máquinas maciçamente paralelas (MPP)

Os MPPs (Massively Parallel Processors) são multicomputadores NORMA construídos com milhares de processadores comerciais conectados por uma rede de alta velocidade. O alto desempenho é obtido com o grande número de processadores. O fato de haver troca de mensagens torna a programação mais difícil que nos casos em que a memória é compartilhada. Exemplos: Intel Paragon (máximo de 4000 processadores), Connection Machine CM-5 (máximo de 2048 processadores), IBM SP2 (máximo de 512 processadores) e Cray T3D (máximo de 2048 processadores).

## Máquinas com memória compartilhada distribuída (DSM)

Nos sistemas DSM (Distributed Shared Memory), mesmo com a memória sendo distribuída entre os nós, todos os processadores podem acessar todas as memórias. O espaço de endereçamento único, o compartilhamento de dados e o controle de coerência de cache são conseguidos com software. Podem ser sistemas NUMA com memória entrelaçada distribuída, ou sistemas NORMA (com memórias locais), onde as memórias podem ser ligadas através de adaptadores de rede (AR) a uma rede de interconexão específica, que permite o acesso a memórias remotas. A máquina, nos dois casos, é considerada CC-NUMA ou SC-NUMA dependendo da implementação da coerência de cache. Exemplo: SGI Origin (máximo de 512 processadores).

# Redes de estações de trabalho (NOW)

As redes de estações de trabalho (NOW – *Network of Workstations*) são constituídas de várias estações de trabalho interligadas por alguma tecnologia tradicional de rede, como Ethernet e ATM. Na prática são redes locais utilizadas na execução de aplicações paralelas. Podem ser vistas como máquinas NORMA de baixo custo, ou sem custo algum caso a rede já exista, ou seja, esta é uma solução significativamente mais barata em relação aos MPPs. A desvantagem clara que se vê em uma rede de estações de trabalho é o fato de que as redes tradicionais costumam ser usadas apenas em tarefas menores (para compartilhar arquivos e acessar impressoras remotas, por exemplo), e geralmente não são otimizadas para operações de comunicação de uma aplicação paralela. O resultado é uma alta latência nessas operações, o que compromete o desempenho da máquina como um todo. São usadas principalmente em instituições de ensino para o estudo de processamento paralelo e distribuído. Exemplo: Estações de trabalho interligadas por tecnologia Ethernet.

# COMPARAÇÃO ENTRE OS MODELOS FÍSICOS PARA CONSTRUÇÃO DE MÁQUINAS PARALELAS

| Tipo | PVP | SMP | MPP | DSM | NOW |
|---|---|---|---|---|---|
| Estrutura | SIMD | MIMD | MIMD | MIMD | MIMD |
| Comunicação | memória compartilhada | memória compartilhada | troca de mensagem | memória compartilhada | troca de mensagem |
| Interconexão | Crossbar | barramento ou crossbar | rede específica | rede específica | rede comum |
| Número de nós | ~10 | ~50 | Acima de 100 | 10 a 2000 | 4 a 10000 |
| Endereçamento | único | único | múltiplo | único | múltiplo |
| Processador | Específico | comum | comum | específico ou comum | comum |
| Acesso à memória | UMA | UMA | NORMA | NUMA | NORMA |

NORMA – Non-remote memory access
UMA – Uniform memory access
NUMA – Non-uniform memory Access

**Algoritmos Paralelos (disciplina regular)**

**Ementa:**

1. Modelos de computação paralela:
   - Memória compartilhada: PRAM síncrona e assíncrona, QSP, QWQR, BSP, LogP, SMP
   - Memória distribuída: SMPD
   - circuitos combinatoriais
2. paradigmas de programação: multi-threaded, BSP, passagem de mensagem, funcional, tuple space, CSP
3. desempenho: profundidade de computação, trabalho, custo, speed-up, eficiência
4. Famílias fundamentais de algoritmos: árvore binária balanceada, divide-and-conquer, jumping pointer, compressão, randomização
5. aplicações
   - ordenação
   - grafos
   - processamento de strings
   - álgebra linear
   - otimização
   - Complexidade paralela: P-completeness

UFES
UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO