



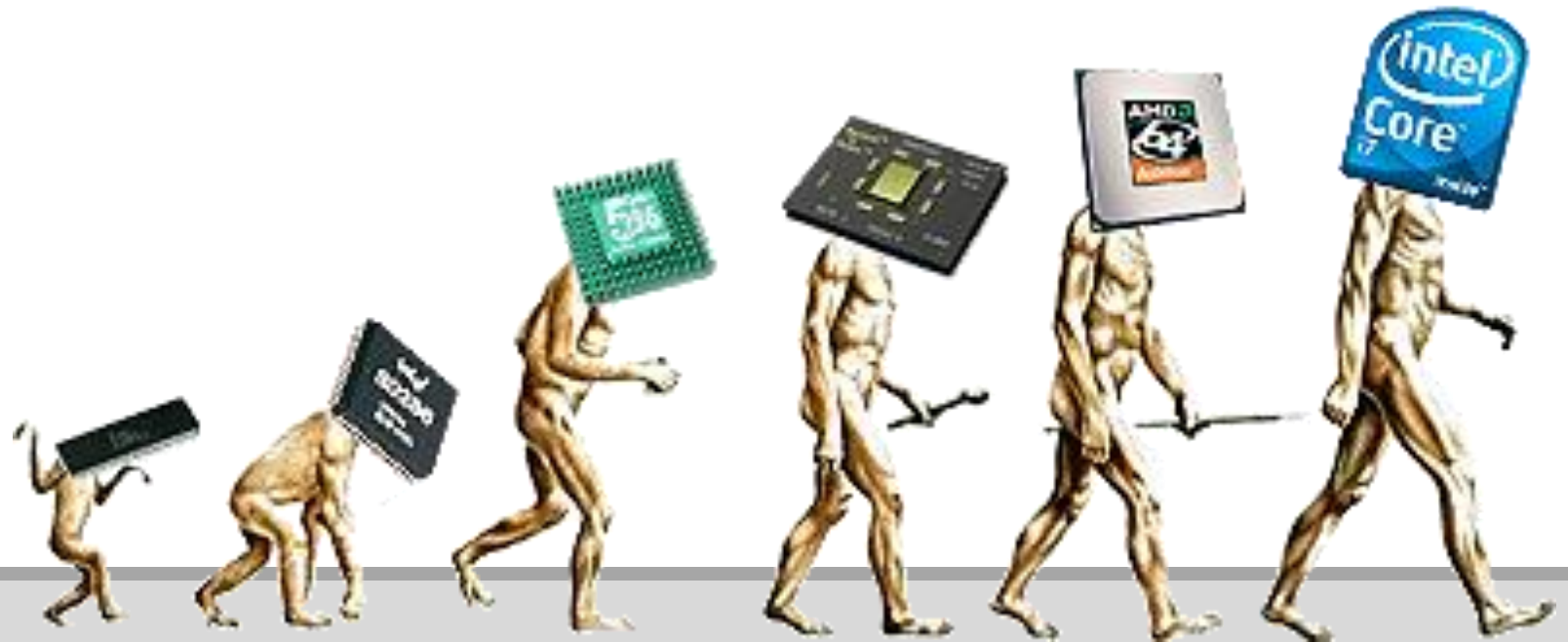
Bacharelado em Sistemas de Informação

Sistemas Operacionais

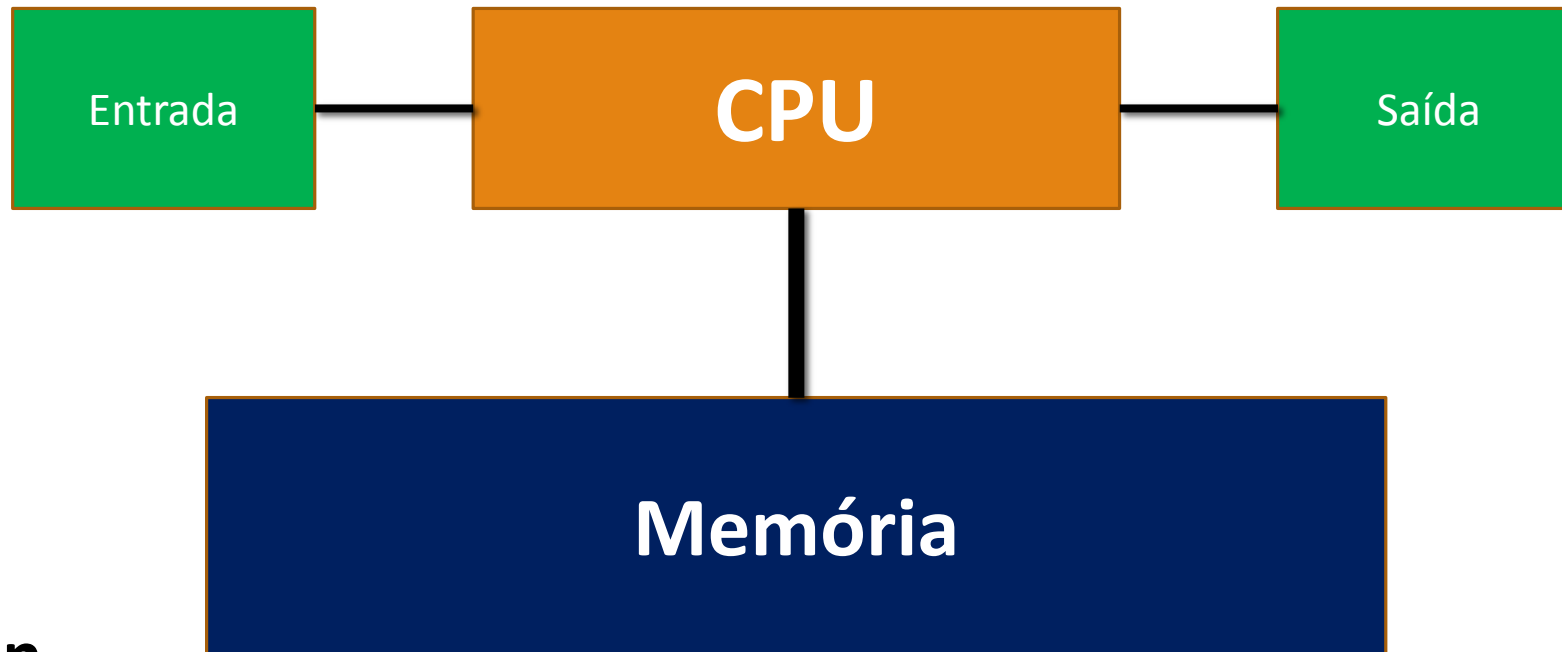
Prof. Filipo Mór



Evolução das Arquiteturas

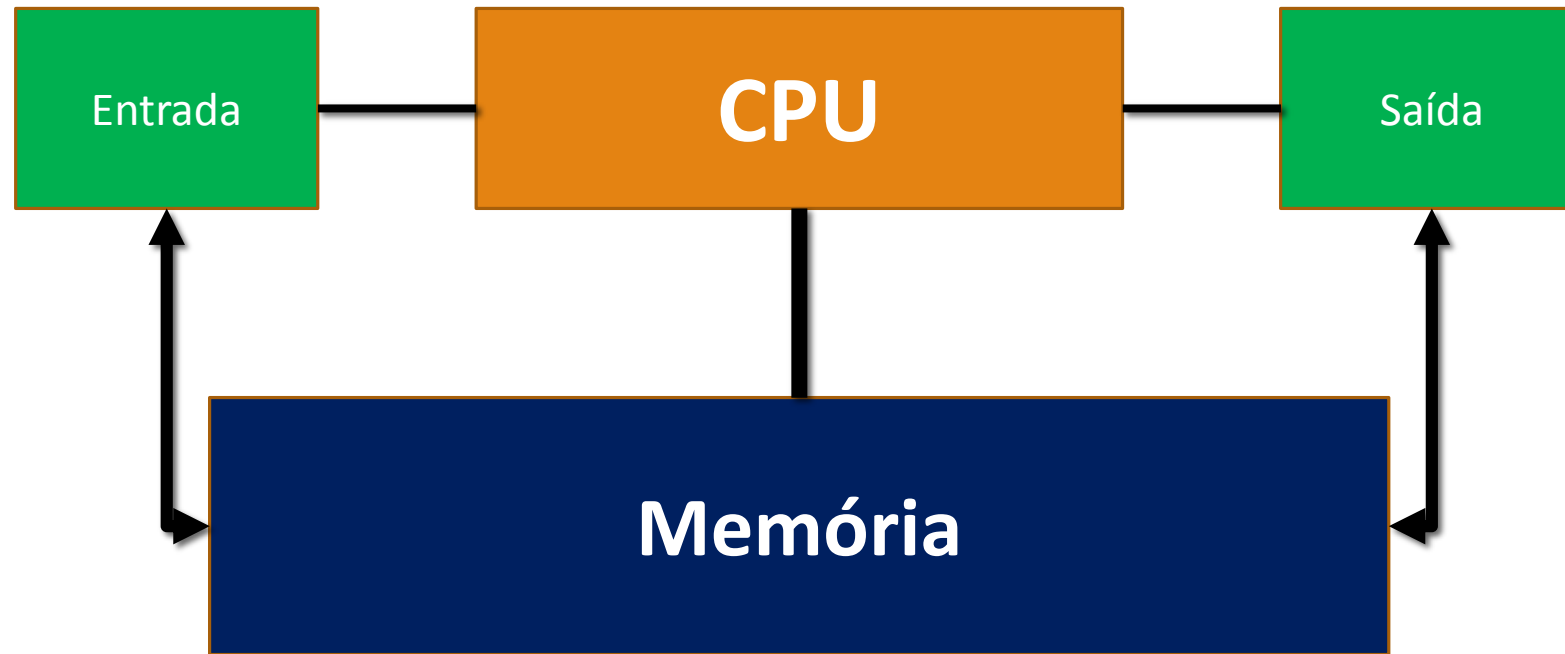


Evolução das Arquiteturas



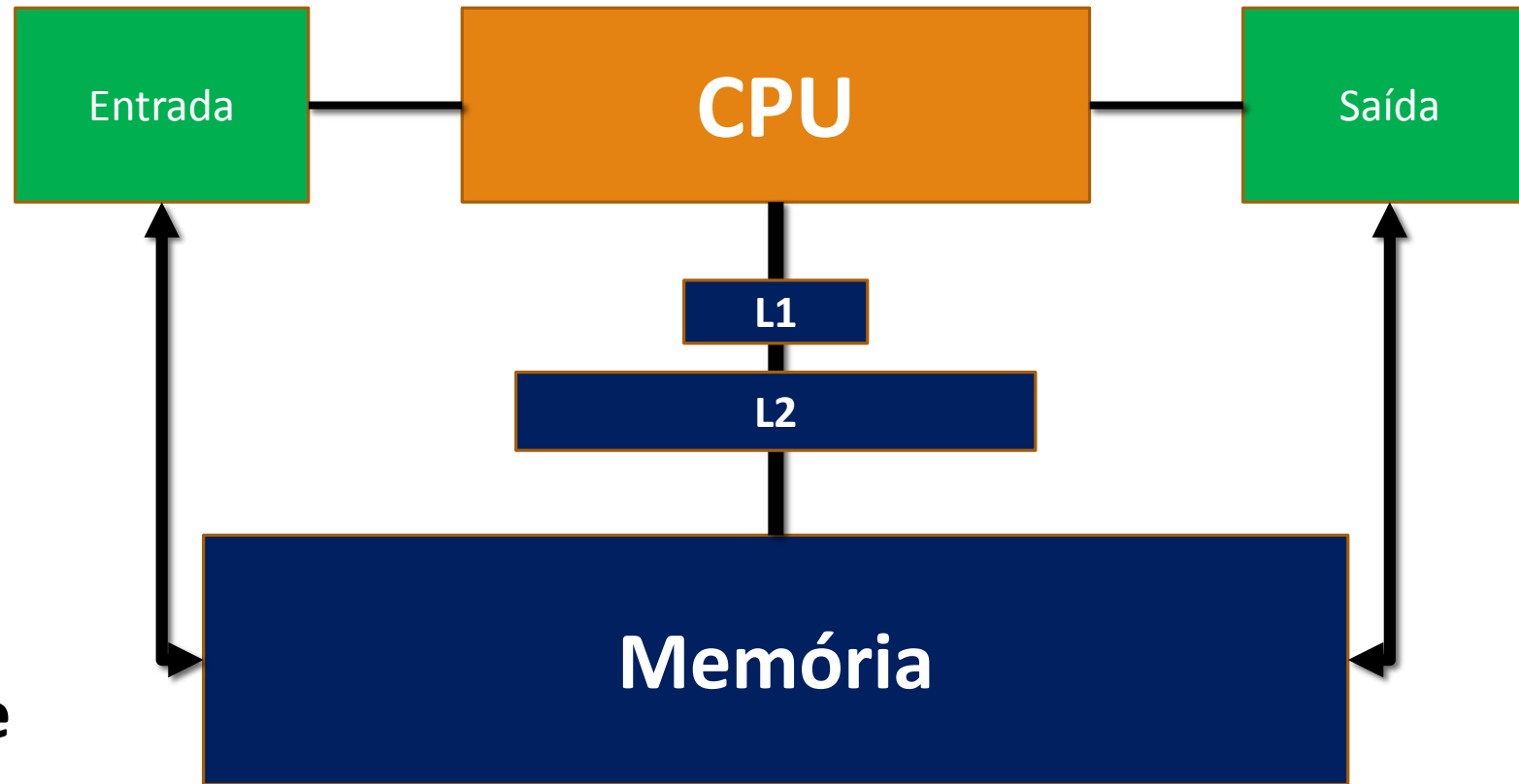
von Neumann

Evolução das Arquiteturas



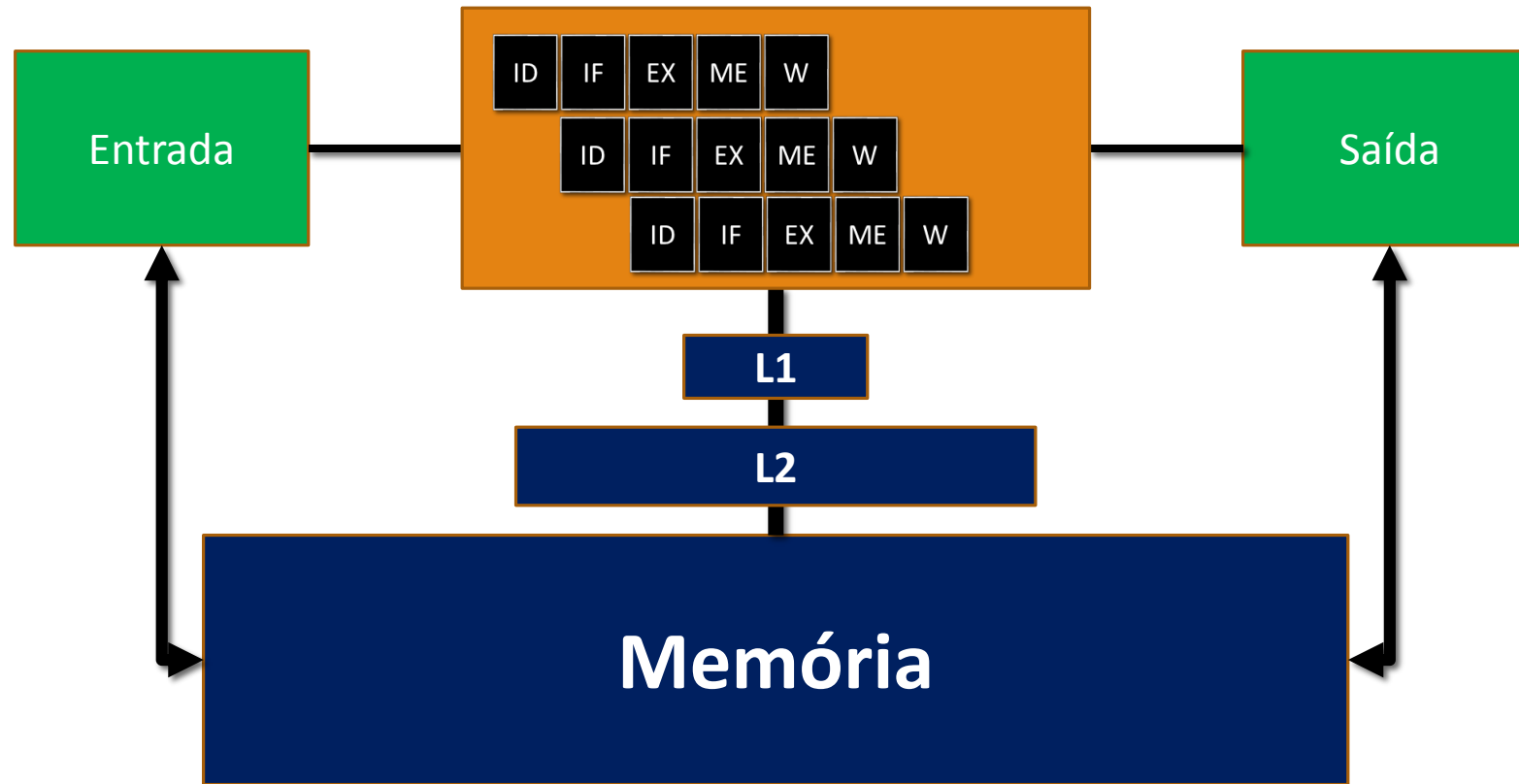
DMA

Evolução das Arquiteturas



Hierarquia de Memória

Evolução das Arquiteturas

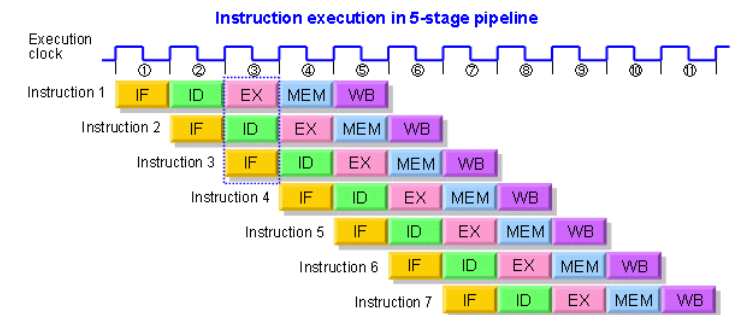


Pipelines:
* Instrução
* Dados

Evolução das Arquiteturas

Objetivo: acelerar o processamento de dados.

- ❑ Delegando tarefas da CPU.
 - ❑ Controle do barramento.
 - ❑ Tratamento de Entrada/Saída.
- ❑ Acelerado o acesso a memória.
 - ❑ Com a hierarquia de memória.
- ❑ Sobrepondo ciclos da CPU.
 - ❑ Com o uso de pipelines.



Replicando a CPU

Opções existentes:

- Pipeline superscalar (*paralelismo de instruções*).
- Máquinas vetoriais.
- Multiprocessadores.
- Multicomputadores.

Replicando a CPU

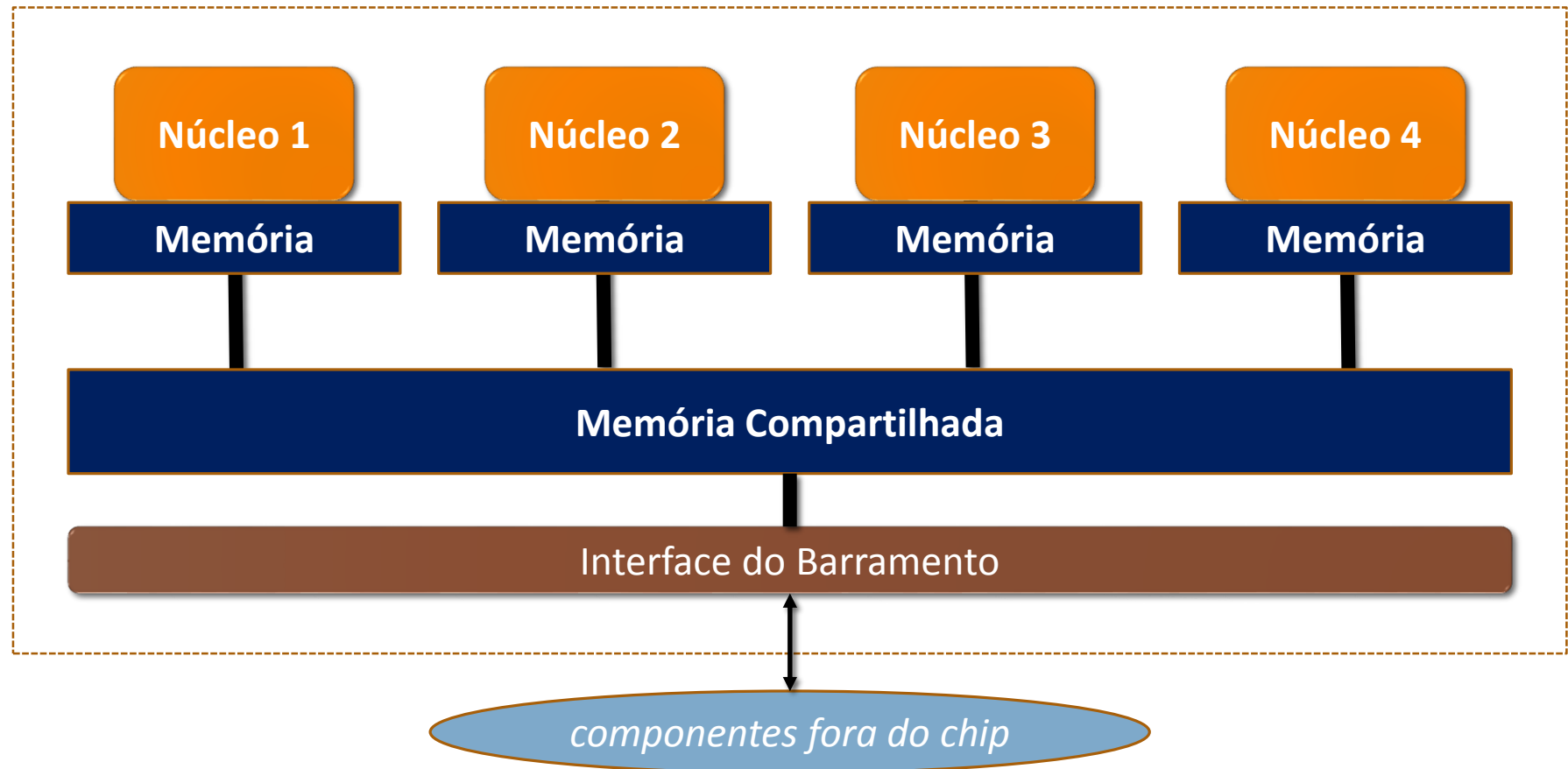


Diagrama de um processador multicore genérico.

Replicando a CPU

Sequencial

1. $A = 3 + 5$
2. $B = 4 + 8$
3. $C = A + B$

3 ciclos
3 instruções

Paralelo

- 1a. $A = 3 + 5$
- 1b. $B = 4 + 8$
2. $C = A + B$

2 ciclos
3 instruções

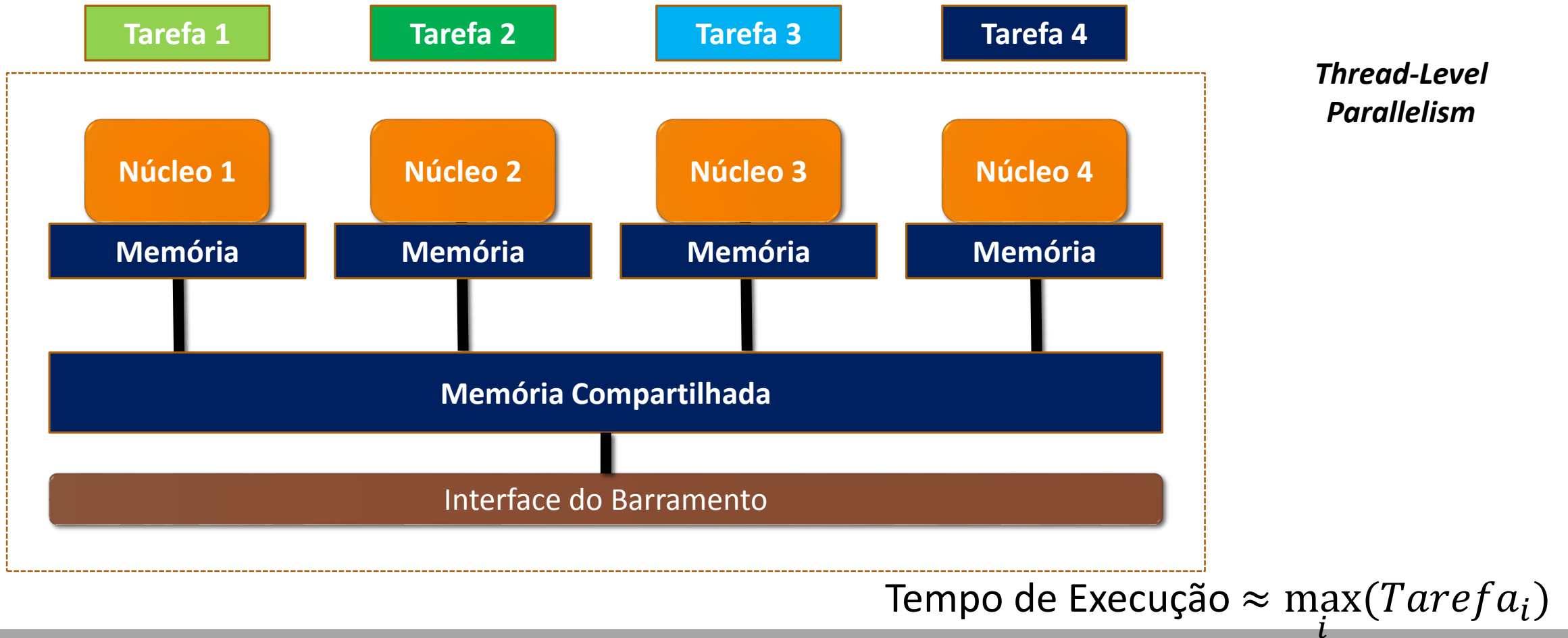
Replicando a CPU

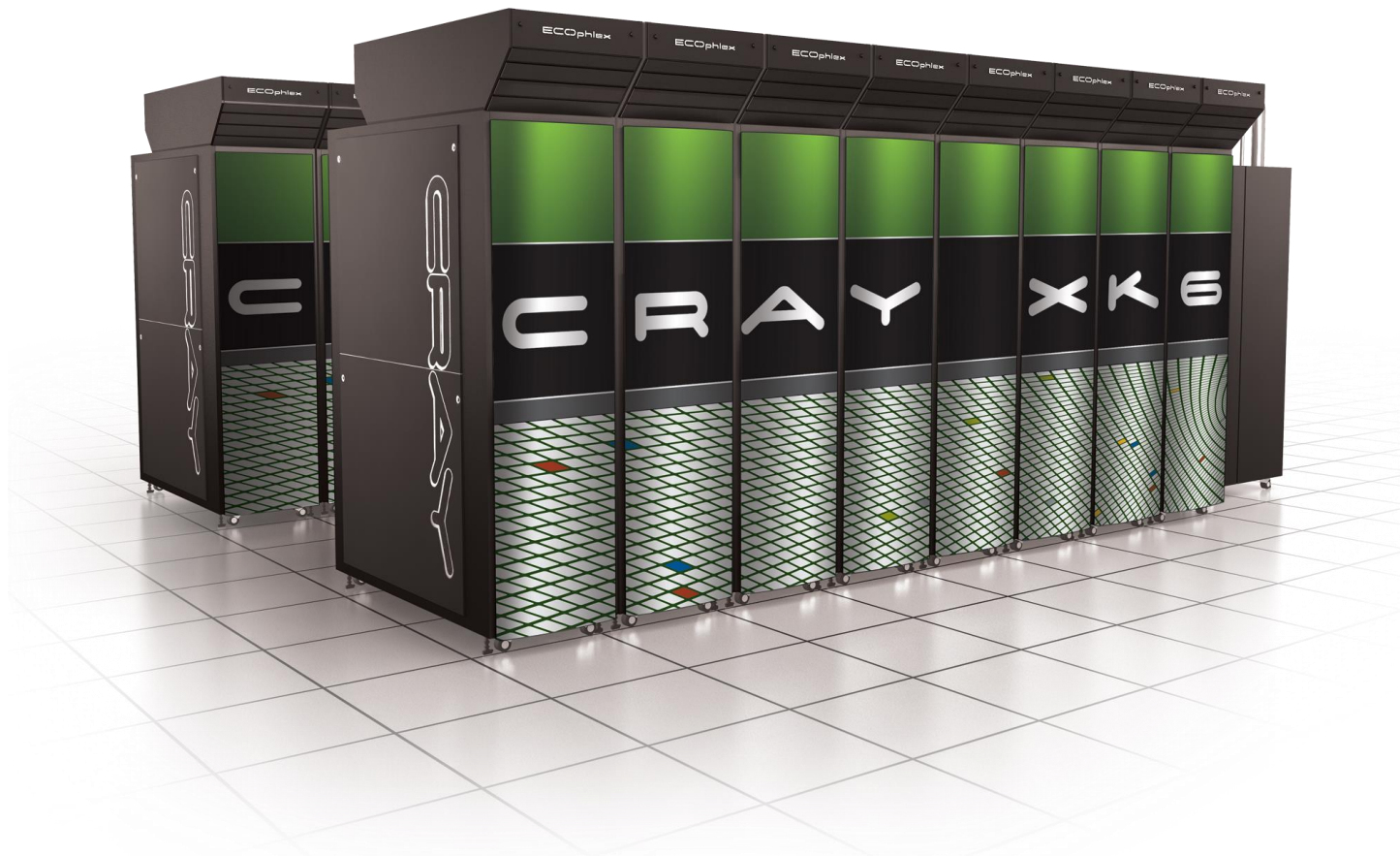
Single Threading



$$\text{Tempo de Execução} \approx \sum_{i=1}^4 Tarefa_i$$

Replicando a CPU





CLASSIFICAÇÃO DE MÁQUINAS PARALELAS

Classificando Máquinas Paralelas

Por que estudar classificações ?

- Para identificar critérios da classificação
 - Por que os critérios são importantes e quais são as suas implicações

- Para analisar todas possibilidades e efeitos destas possibilidades nas arquiteturas
 - Mesmo para classes que não foram implementadas ou implementações que não deram certo

- Para entender como se deu a evolução da área

- Para planejar a evolução da área

Classificando Máquinas Paralelas

Classificação de Flynn

- Estabelecida na década de 1970.
- Diferencia as máquinas de acordo com o fluxo de instruções e o fluxo de dados.

Classificando Máquinas Paralelas

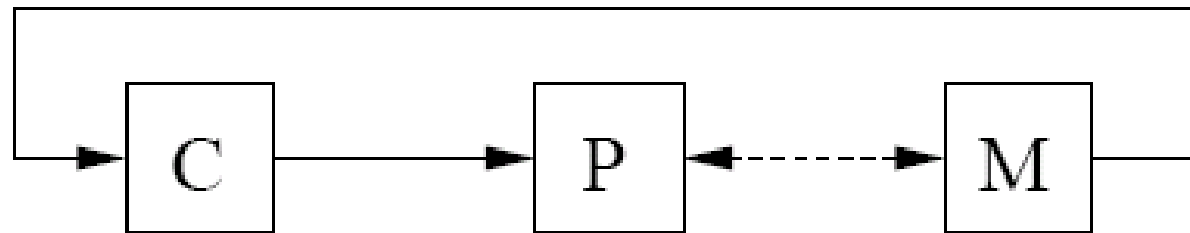
	Single Data	Multiple Data
Single Instruction	SISD	SIMD
Multiple Instructions	MISD	MIMD

Classificação de Flynn

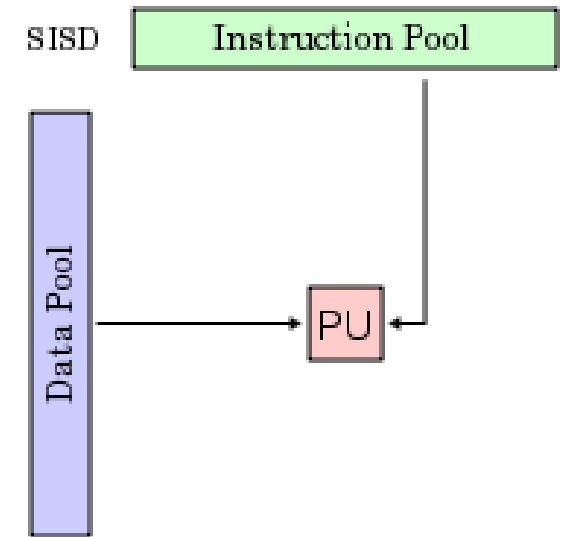
Classificando Máquinas Paralelas

Classe SISD

- ❑ Fluxo de instruções único
- ❑ Fluxo de dados único
- ❑ Arquiteturas tradicionais Von Neumann, primeiros computadores pessoais e estações de trabalho.



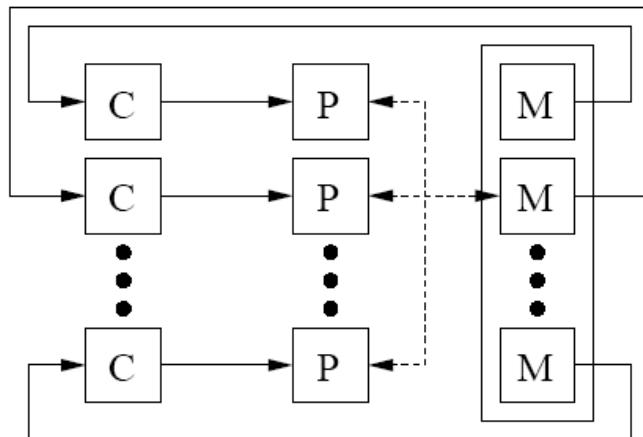
—→ Instruções
←---→ Dados



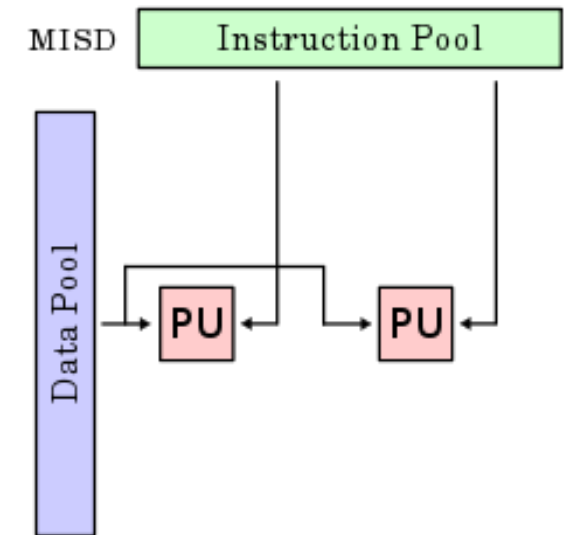
Classificando Máquinas Paralelas

Classe MISD

- ❑ Múltiplos Fluxos de instruções.
- ❑ Fluxo de dados único.
- ❑ Nenhuma máquina implementada até o momento.



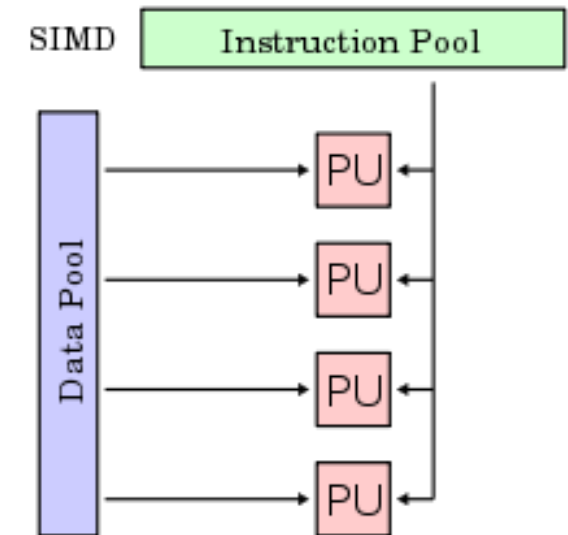
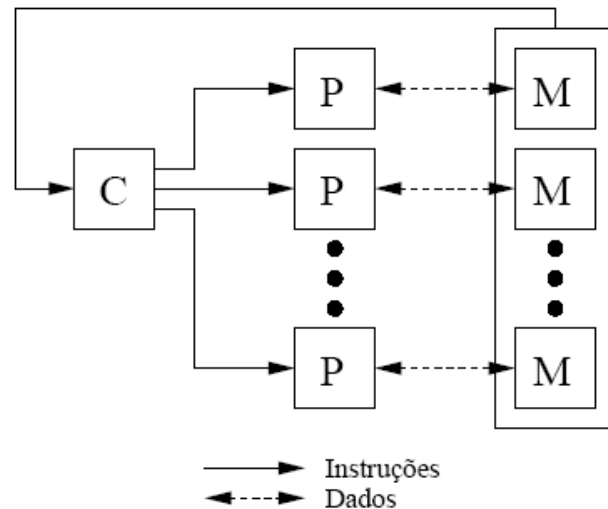
→ Instruções
←--- Dados



Classificando Máquinas Paralelas

Classe SIMD

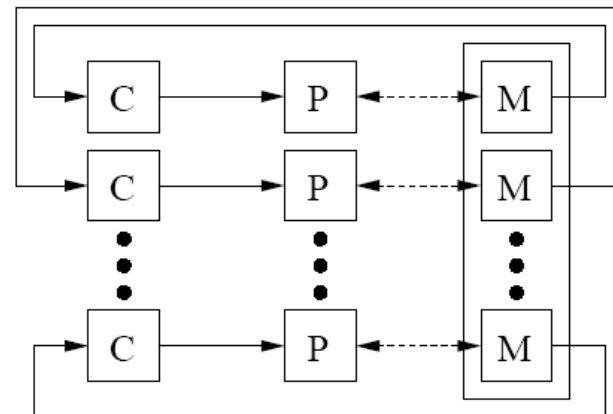
- ❑ Fluxos de instruções Único.
- ❑ Múltiplos Fluxos de dados.
- ❑ Execução Síncrona.
- ❑ GPUs, arquiteturas em Array.



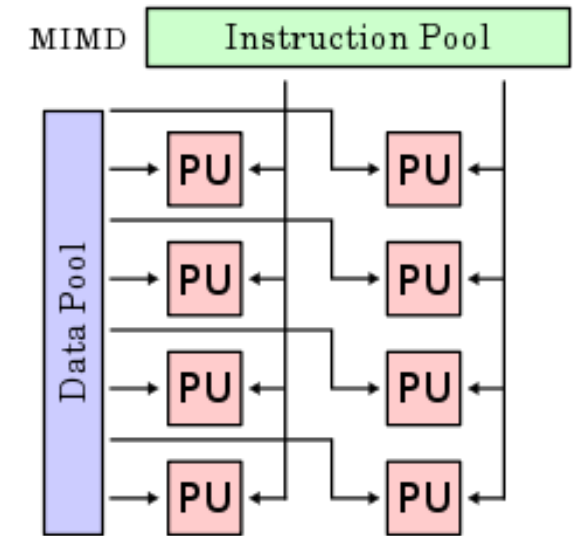
Classificando Máquinas Paralelas

Classe MIMD

- ❑ Múltiplos Fluxos de instruções.
- ❑ Múltiplos Fluxos de dados.
- ❑ Arquiteturas modernas.
- ❑ GPUs mais recentes, aceleradores *manycore* (Intel Xion Phi).



→ Instruções
←- - - Dados



Distribuição de Memória

Refere-se à localização física da memória

Memória distribuída (*distributed memory*)

- Memória implementada com vários módulos
- Cada módulo fica próximo de um processador

Memória centralizada (*centralized memory*)

- Memória encontra-se à mesma distância de todos os processadores → Independentemente de ter sido implementada com um ou vários módulos

Compartilhamento de Memória

Refere-se ao espaço de endereçamento dos processadores

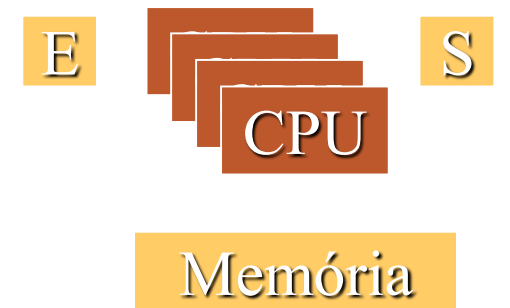
Memória compartilhada (shared memory)

- Único espaço de endereçamento usado para comunicação entre processadores
- Operações de load e store

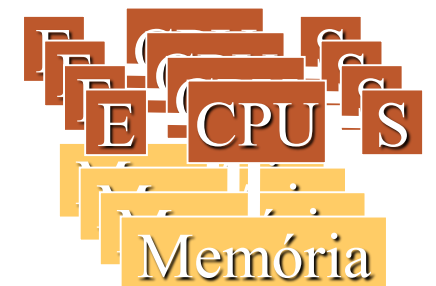
Memória não compartilhada

- Múltiplos espaços de endereçamento privados (*multiple private address spaces*) → um para cada processador
- Comunicação através de troca de mensagens → operações send e receive

Conforme compartilhamento de memória as máquinas podem ser:



Multiprocessadores



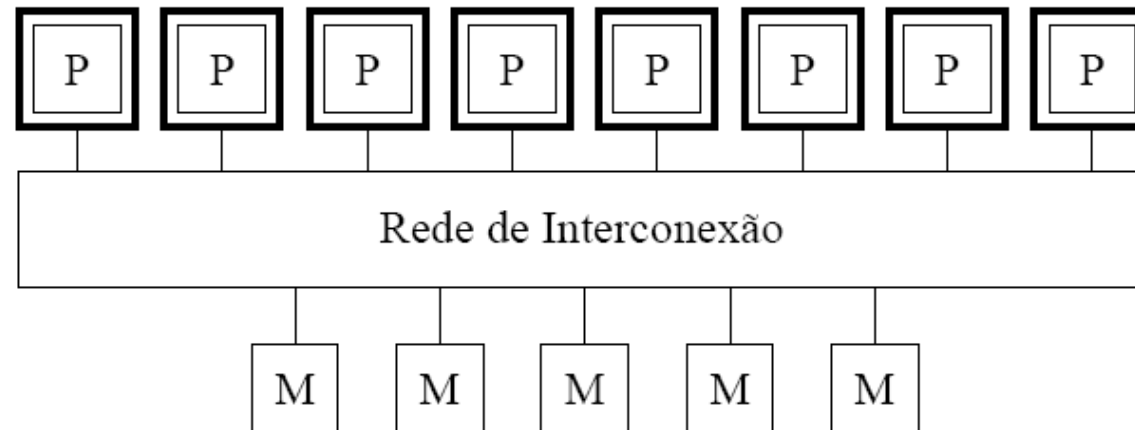
Multicomputadores

Multiprocessadores

Máquina paralela construída a partir da replicação de processadores de uma arquitetura convencional

Todos os processadores (P) acessam memórias compartilhadas (M) através de uma infra-estrutura de comunicação

- Possui apenas um espaço de endereçamento
- Comunicação entre processos de forma bastante eficiente (load e store)



Em relação ao tipo de acesso às memórias do sistema, multiprocessadores podem ser classificados como

- UMA
- NUMA, NCC-NUMA, CC-NUMA
- COMA

Acesso Uniforme à Memória

(Uniform Memory Access - UMA)

Memória centralizada

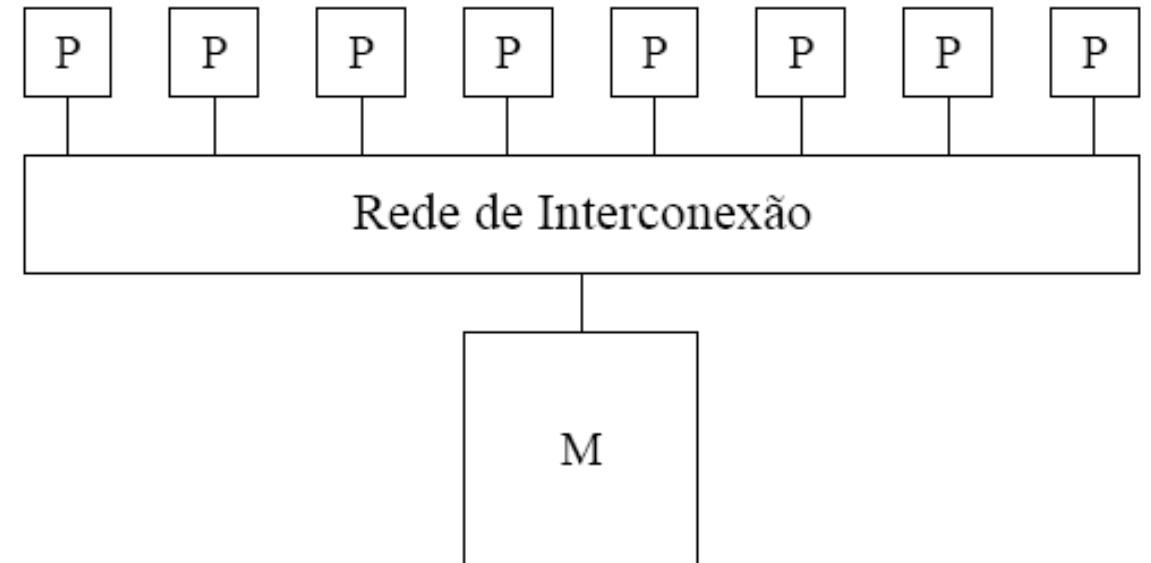
- Encontra-se à mesma distância de todos processadores

Latência de acesso à memória

- Igual para todos processadores

Infra-estrutura de comunicação

- Barramento é a mais usada → suporta apenas uma transação por vez
- Outras infra-estruturas também se enquadram nesta categoria, se mantiverem uniforme o tempo de acesso à memória



Acesso Não Uniforme à Memória

(*Non-Uniform Memory Access* - NUMA)

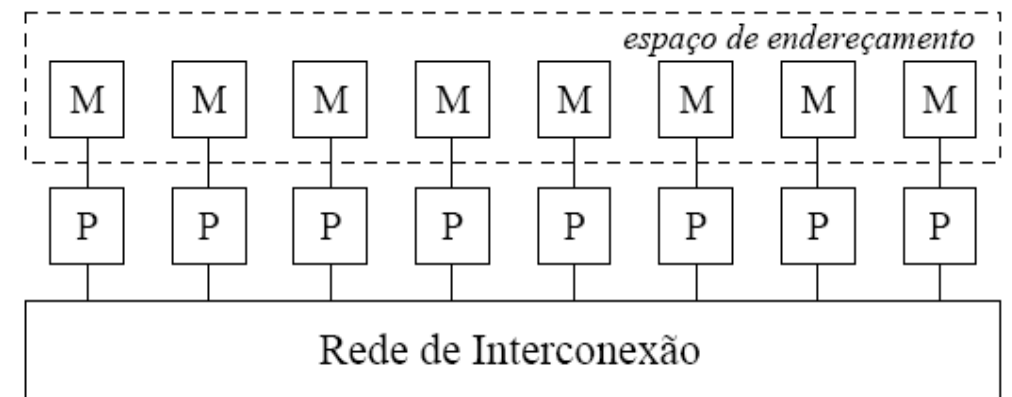
Memória Distribuída

Espaço de endereçamento único

Memória implementada com múltiplos módulos associados a cada processador

Comunicação processador-memórias não locais através da infra-estrutura de comunicação

Tempo de acesso à memória local < tempo de acesso às demais → **Acesso não uniforme** → Distância das memórias variável → depende do endereço



Arquiteturas de Memória Somente com *Cache*

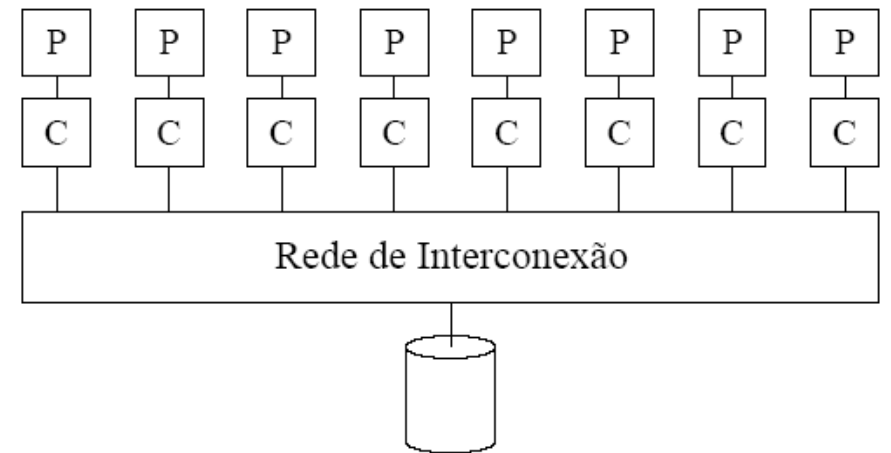
(*Cache-Only Memory Architecture* - COMA)

Memórias locais estão estruturadas como memórias *cache*

- São chamadas de COMA *caches*
- COMA *caches* têm muito mais capacidade que uma *cache* tradicional

Arquiteturas COMA têm suporte de hardware para a replicação efetiva do mesmo bloco de *cache* em múltiplos nós

- Reduz tempo global para pegar informações



Multicomputadores

Processadores (P) possuem memória local (M) de acesso restrito

Memórias de outros processadores são remotas → Possuem espaços de endereçamento distintos

Não é possível uso de variáveis compartilhadas

- Troca de informações feita por envio de mensagens
- Máquinas também chamadas *message passing systems*

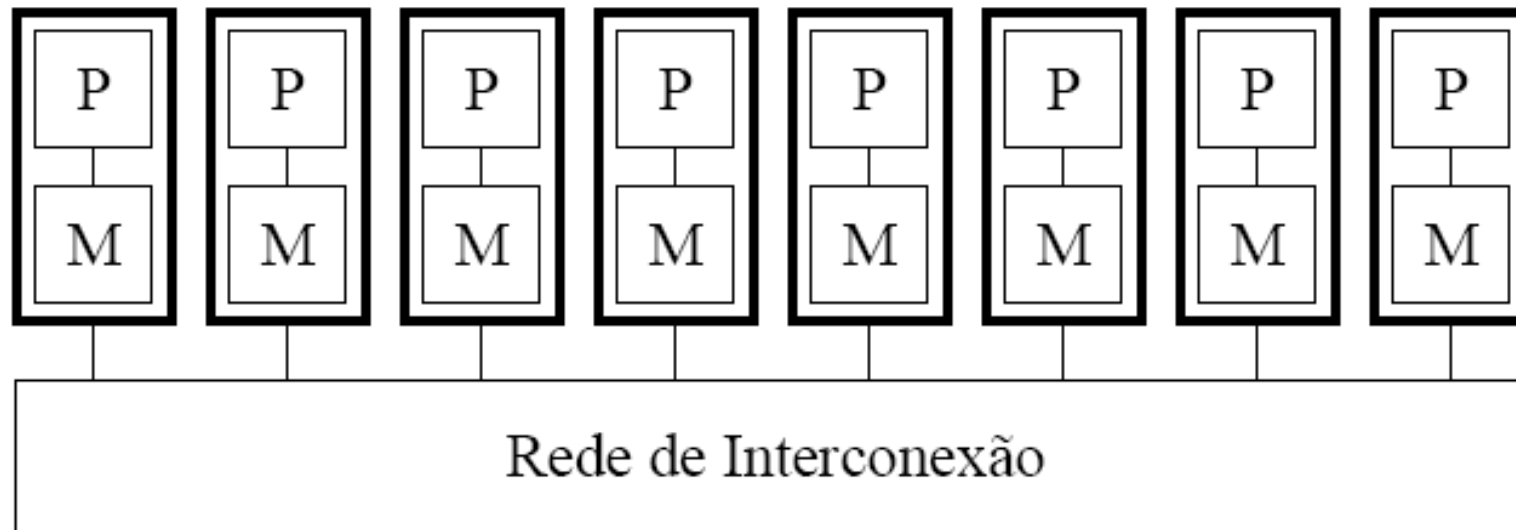
Multicomputadores são construídos a partir da replicação de toda arquitetura convencional, não apenas do processador, como nos **multiprocessadores**

Sem Acesso à Memória Remota

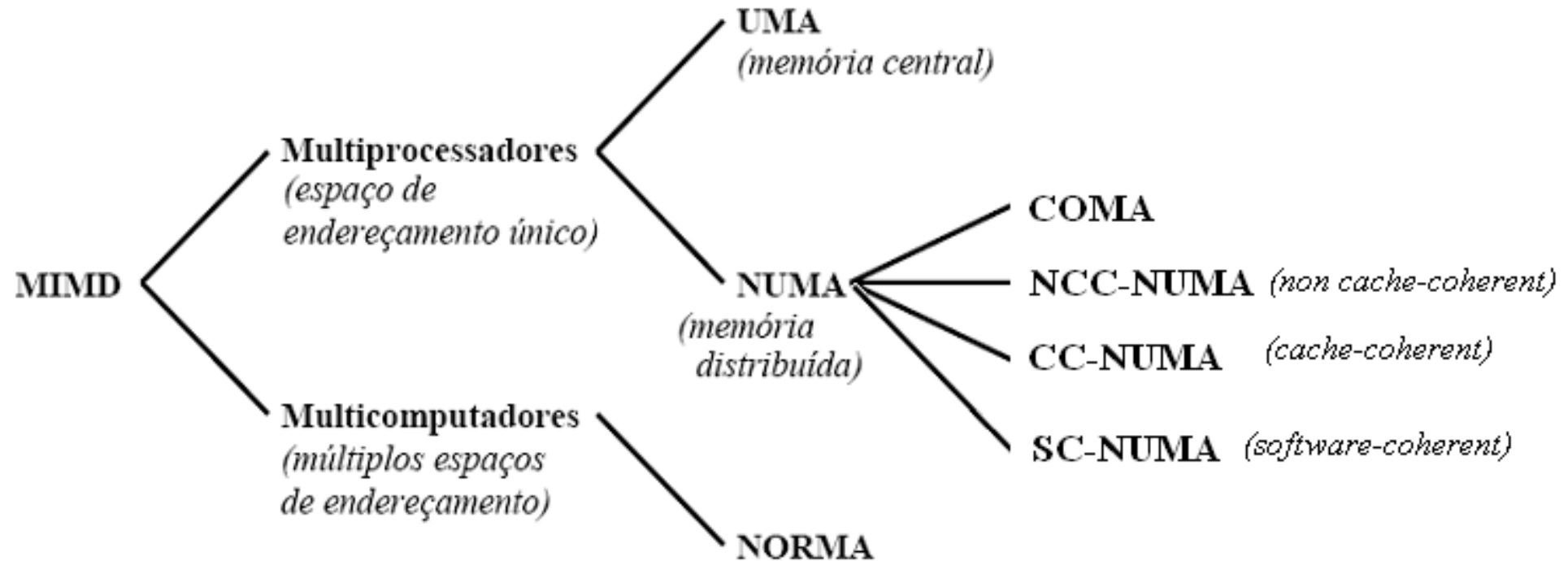
(Non-Remote Memory Access - NORMA)

Multicomputadores são classificados como NORMA

Replicação inteira da arquitetura faz que cada nó só consiga endereçar sua memória local



Visão Geral da Classificação Segundo Modelo de Memória





EXERCÍCIOS

Exercícios

1. Quais são as duas possibilidades para se construir máquinas com vários processadores? Que nome é dado a cada uma delas?
2. Qual o efeito prático de ter um espaço de endereçamento único? E um espaço de endereçamento distinto?
3. Discuta a classificação de máquinas paralelas segundo a memória (compartilhamento/distribuição)
4. Qual a principal diferença entre memórias centralizadas e memórias distribuídas? O que caracteriza esta classificação?
5. Discuta a afirmação: - “A programação paralela com memória de endereçamento compartilhado é mais simples que a programação com endereçamento distribuído”
6. Compare o desempenho de multicomputadores com multiprocessadores
7. Quais as vantagens das máquinas UMA/NUMA sobre as máquinas NORMA, sob o ponto de vista do programador? Por quê?
8. Diferencie as arquiteturas com organização de memória tipo UMA/NUMA das arquiteturas tipo NORMA

Exercícios

9. Por que alguns autores consideram uma arquitetura COMA como subclassificação de arquiteturas NUMA?
10. Qual classe de máquinas possui uma memória distribuída e compartilhada?
11. Discuta a afirmação: - “Uma máquina COMA tem maior desempenho que uma máquina NORMA”. Diga se verdadeira ou não e quando?
12. Qual o efeito de trabalhar em máquinas NUMA, se comparados com máquinas UMA?
13. Desenhe um esquema básico de arquiteturas NORMA e NUMA. Os desenhos são muito parecidos. Onde está a diferença significativa entre eles?
14. Como a máquina UMA, que se utiliza de um barramento para ligar os processadores à memória principal se beneficiaria dos múltiplos canais de uma memória entrelaçada?
15. Qual a grande vantagem de uma máquina COMA frente a uma máquina NUMA?
16. Quais são as características das arquiteturas UMA (acesso uniforme à memória) e NUMA (acesso não uniforme à memória), quanto: (a) número de processadores, (b) rede de interconexão, (c) tempo de acesso à memória? Faça um esquema em blocos das arquiteturas UMA e NUMA

Exercícios

17. (ENADE 2008 - adaptado) Uma alternativa para o aumento de desempenho é o uso de processadores com múltiplos núcleos (multicores). Nesses sistemas, cada núcleo, normalmente, tem funcionalidades completas de um processador. Atualmente é comum configurações com 4 ou mais núcleos. Com relação ao uso de multicores, e sabendo que *threads* são estruturas de execução associadas a um processo, que compartilham áreas de código e dados, mantendo contextos independentes, analise as seguintes asserções:

Tendo suas atividades divididas em múltiplos threads que podem ser executadas paralelamente, aplicações podem se beneficiar mais efetivamente dos diversos núcleos dos processadores multicores

porque

o SO que executa em processadores multicores pode alocar os núcleos existentes para executar simultaneamente diversas sequências de código, sobrepondo suas execuções e, normalmente, reduzindo o tempo de resposta das aplicações às quais estão associadas.

Acerca dessas asserções, assinale a opção correta:

- i. As duas asserções são proposições verdadeiras, sendo a segunda uma justificativa correta da primeira
- ii. As duas asserções são proposições verdadeiras, mas a segunda não é uma justificativa correta da primeira.
- iii. A primeira asserção é uma proposição verdadeira, e a segunda, uma proposição falsa.
- iv. A primeira asserção é uma proposição falsa, e a segunda, uma proposição verdadeira.
- v. Tanto a primeira quanto a segunda asserções são proposições falsas.

Exercícios

17. (ENADE 2008 - adaptado) Uma alternativa para o aumento de desempenho é o uso de processadores com múltiplos núcleos (multicores). Nesses sistemas, cada núcleo, normalmente, tem funcionalidades completas de um processador. Atualmente é comum configurações com 4 ou mais núcleos. Com relação ao uso de multicores, e sabendo que *threads* são estruturas de execução associadas a um processo, que compartilham áreas de código e dados, mantendo contextos independentes, analise as seguintes asserções:

Tendo suas atividades divididas em múltiplos threads que podem ser executadas paralelamente, aplicações podem se beneficiar mais efetivamente dos diversos núcleos dos processadores multicores

porque

o SO que executa em processadores multicores pode alocar os núcleos existentes para executar simultaneamente diversas sequências de código, sobrepondo suas execuções e, normalmente, reduzindo o tempo de resposta das aplicações às quais estão associadas.

Acerca dessas asserções, assinale a opção correta:

- i. As duas asserções são proposições verdadeiras, sendo a segunda uma justificativa correta da primeira**
- ii. As duas asserções são proposições verdadeiras, mas a segunda não é uma justificativa correta da primeira.
- iii. A primeira asserção é uma proposição verdadeira, e a segunda, uma proposição falsa.
- iv. A primeira asserção é uma proposição falsa, e a segunda, uma proposição verdadeira.
- v. Tanto a primeira quanto a segunda asserções são proposições falsas.

Exercícios

18. (POSCOMP 2011, Questão 29) Relacione a coluna da esquerda com a coluna da direita

(I) Multicore	(A) Múltiplos pipelines que operam em paralelo
(II) Superpipeline	(B) Execução de instruções fora de ordem em um pipeline
(III) Superescalar	(C) Pipelines com grande número de estágios
(IV) Pipeline dinâmico	(D) Múltiplos processadores compartilhando um espaço de endereços
(V) Multiprocessadores	(E) Múltiplos processadores em um único encapsulamento

Assinale a alternativa que contém a associação correta

- a) I-B, II-A, III-C, IV-E, V-D
- b) I-C, II-A, III-B, IV-D, V-E
- c) I-D, II-E, III-B, IV-A, V-C
- d) I-E, II-C, III-A, IV-B; V-D
- e) I-E, II-C, III-A, IV-D, V-B

Exercícios

18. (POSCOMP 2011, Questão 29) Relacione a coluna da esquerda com a coluna da direita

(I) Multicore	(A) Múltiplos pipelines que operam em paralelo
(II) Superpipeline	(B) Execução de instruções fora de ordem em um pipeline
(III) Superescalar	(C) Pipelines com grande número de estágios
(IV) Pipeline dinâmico	(D) Múltiplos processadores compartilhando um espaço de endereços
(V) Multiprocessadores	(E) Múltiplos processadores em um único encapsulamento

Assinale a alternativa que contém a associação correta

a) I-B, II-A, III-C, IV-E, V-D

b) I-C, II-A, III-B, IV-D, V-E

c) I-D, II-E, III-B, IV-A, V-C

d) I-E, II-C, III-A, IV-B; V-D

e) I-E, II-C, III-A, IV-D, V-B

Exercícios

19. (POSCOMP 2011, Questão 41) O gerenciamento de processos em sistemas modernos é feito, quase sempre, com o uso de preempção de processos através de técnicas de compartilhamento de tempo. O que a introdução de processadores com vários núcleos altera nesse gerenciamento?
- a) Torna-se possível a paralelização efetiva de processos concorrentes.
 - b) Torna-se possível eliminar a condição de corrida em processos concorrentes executados em paralelo
 - c) Torna-se possível o uso de threads para a execução de processos concorrentes
 - d) Torna-se possível separar os demais mecanismos de gerenciamento do sistema operacional do gerenciamento de processos
 - e) Torna-se possível o uso de sistemas operacionais multitarefas

Exercícios

19. (POSCOMP 2011, Questão 41) O gerenciamento de processos em sistemas modernos é feito, quase sempre, com o uso de preempção de processos através de técnicas de compartilhamento de tempo. O que a introdução de processadores com vários núcleos altera nesse gerenciamento?

a) Torna-se possível a paralelização efetiva de processos concorrentes.

b) Torna-se possível eliminar a condição de corrida em processos concorrentes executados em paralelo

c) Torna-se possível o uso de threads para a execução de processos concorrentes

d) Torna-se possível separar os demais mecanismos de gerenciamento do sistema operacional do gerenciamento de processos

e) Torna-se possível o uso de sistemas operacionais multitarefas

Exercícios

20. (POSCOMP 2012, Questão 46) Com relação a processadores, considere as afirmativas a seguir

I. Arquiteturas Superescalares podem executar instruções concorrentemente em pipelines diferentes

II. O superpipeline permite a execução de duas tarefas em um único ciclo de clock do processador

III. Multiprocessadores simétricos compartilham a utilização da memória principal

IV. A utilização de uma memória cache L2 compartilhada em processadores multicore é vantajosa em threads que possuem alta localidade

Assinale a alternativa correta:

a) Somente as afirmativas I e II são corretas

b) Somente as afirmativas I e IV são corretas

c) Somente as afirmativas III e IV são corretas

d) Somente as afirmativas I, II e III são corretas

e) Somente as afirmativas II, III e IV são corretas

Exercícios

20. (POSCOMP 2012, Questão 46) Com relação a processadores, considere as afirmativas a seguir

I. Arquiteturas Superescalares podem executar instruções concorrentemente em pipelines diferentes

II. O superpipeline permite a execução de duas tarefas em um único ciclo de clock do processador

III. Multiprocessadores simétricos compartilham a utilização da memória principal

IV. A utilização de uma memória cache L2 compartilhada em processadores multicore é vantajosa em threads que possuem alta localidade

Assinale a alternativa correta:

a) Somente as afirmativas I e II são corretas

b) Somente as afirmativas I e IV são corretas

c) Somente as afirmativas III e IV são corretas

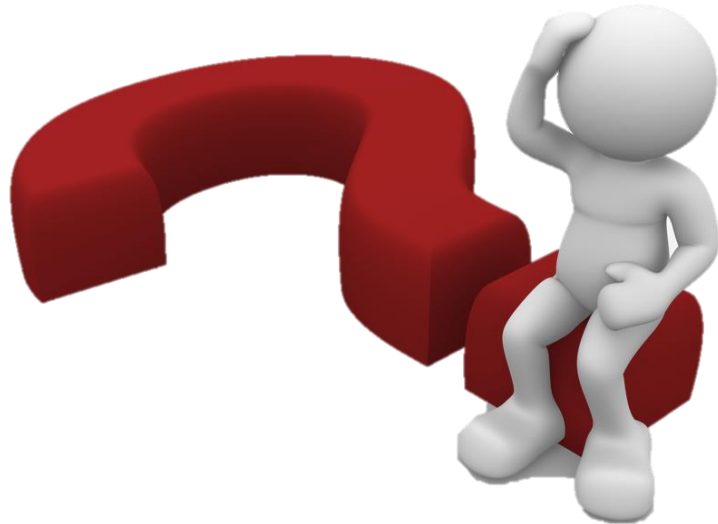
d) Somente as afirmativas I, II e III são corretas

e) Somente as afirmativas II, III e IV são corretas

Bacharelado em Sistemas de Informação

Sistemas Operacionais

Prof. Filipo Mór



Agradecimento especial ao Prof. César Augusto Missio Marcon pelo material gentilmente cedido.

DÚVIDAS?

