

Processamento Paralelo

por

Helcio Wagner da Silva

Introdução

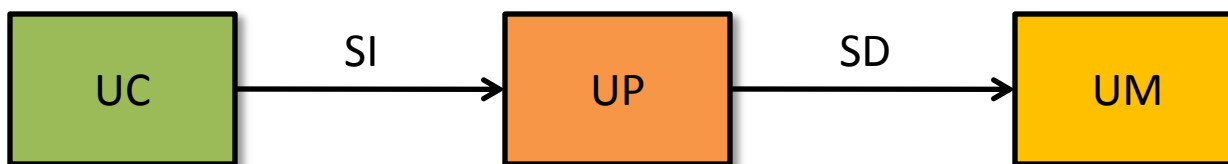
- Tradicionalmente, o computador tem sido visto como uma máquina seqüencial
- Esta visão nunca foi completamente verdadeira
 - No nível das μO , vários sinais de controle são gerados ao mesmo tempo
 - A técnica de pipeline sobrepõe instruções, pelo menos durante as operações de busca e execução
 - Em máquinas superescalares, várias instruções podem ser executadas paralelamente

Introdução

- Novas alternativas para o paralelismo
 - Multiprocessadores simétricos (SMP)
 - Clusters
 - Maquinas com Acesso Não-uniforme a Memória (NUMA)
- Surgiram a partir de novas tecnologias e do barateamento do HW
- Objetivos
 1. Melhoria do desempenho
 2. Aumento da disponibilidade (tolerância a falhas)

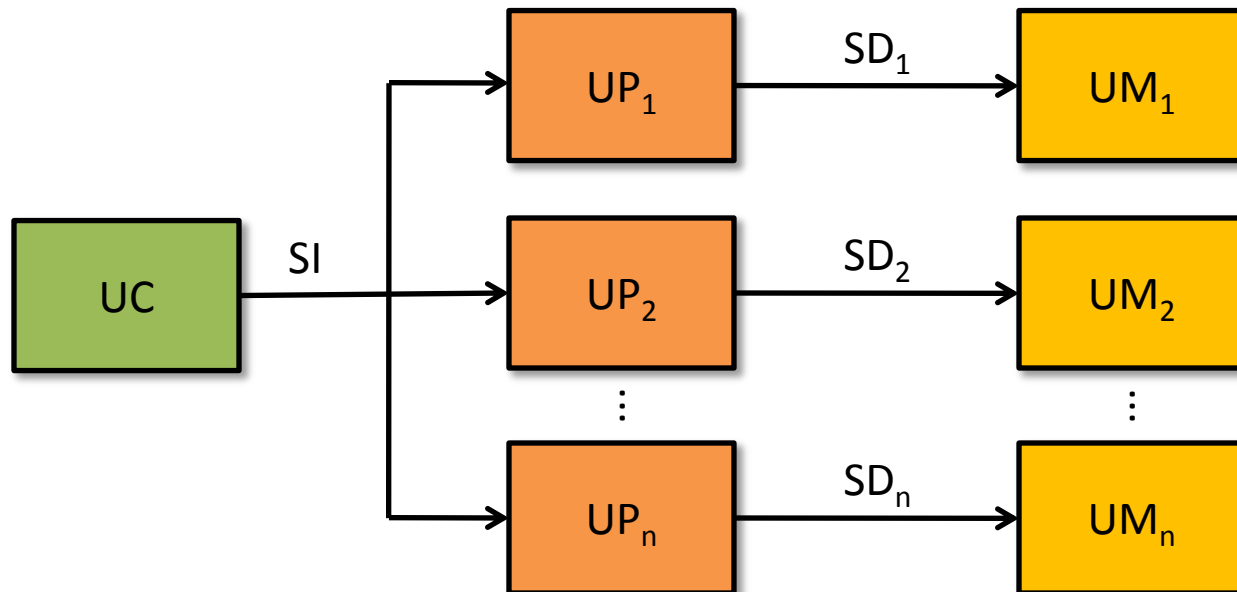
Organizações de Múltiplos Processadores

- SISD (Única Instrução, Único Dado)
 - Unidade de Controle (UC) fornece uma seqüência de instruções para a Unidade de Processamento (UP)
 - UP opera sobre uma única Seqüência de Dados (SD) de uma única Unidade de Memória (UM)
 - Exemplo: sistema uniprocessado



Organizações de Múltiplos Processadores

- SIMD (Única Instrução, Múltiplos Dados)
 - UC alimenta várias UP
 - Cada UP opera sobre uma única SD de uma UM
 - Exemplo: Processadores vetoriais e matriciais



Arquiteturas SIMD

Processadores Vetoriais

Processadores Vetoriais provêm instruções de alto nível sobre vetores de dados, tais como multiplicar, subtrair, somar

Em máquinas escalares, essas operações são realizadas através de um *loop*

Em máquinas vetoriais, essas operações são realizadas em uma única instrução vetorial

Arquiteturas SIMD

Processadores Vetoriais

Vantagens:

- Redução da quantidade de *fetch* e *decode* de instruções.
- Não há necessidade de verificação de conflitos de dados, pois as operações entre elementos dos vetores que estão na mesma operação são independentes.
- Operações vetoriais são atômicas e eliminam a sobrecarga gerada pelos saltos condicionais e comparações necessárias ao controle de repetições.
- As operações entre elementos podem ser paralelizadas ou executadas em pipeline.
- Como há uma quantidade menor de instruções por programa, há uma quantidade menor de falha no *cache* de instruções.

Arquiteturas SIMD

Processadores Vetoriais

Os processadores vetoriais podem ser de dois tipos:

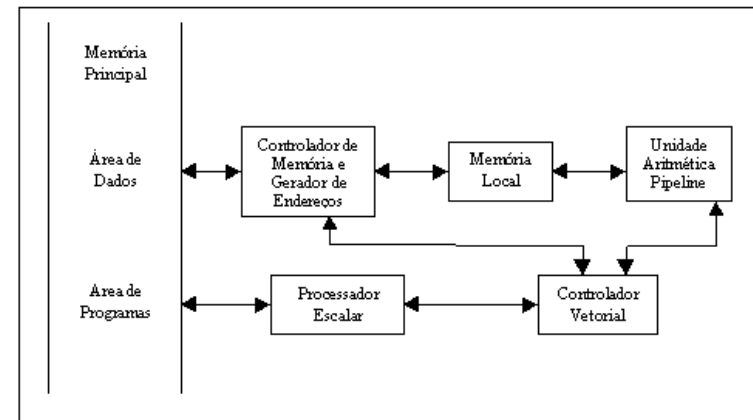
- Registrador-Registrador
 - Operações vetoriais sobre registradores
 - Somente as operações de *load vector* e *store vector* acessam a memória
- Memória-Memória
 - Operandos buscados diretamente da memória

Arquiteturas SIMD

Processadores Vetoriais

Arquitetura de um processador vetorial genérico

- Banco de registradores vetoriais
- Gerador de endereço de vetor:
 - Responsável por gerar o endereço de um vetor a ser buscado na memória.
- Memória
- Unidades aritméticas *pipeline* (*Functional units*):
 - Efetuam as operações sobre os vetores.
- Controlador vetorial (*Vector Data Switch*):
 - Responsável por direcionar um vetor para a unidade



Arquiteturas SIMD

Processadores Vetoriais

Processador Escalar:

- Realiza operações sequenciais, como por exemplo, testes de dependências, gerenciamento de E/S e memória, etc
- Quando identifica uma operação vetorial, encaminha para o Controlador Vetorial

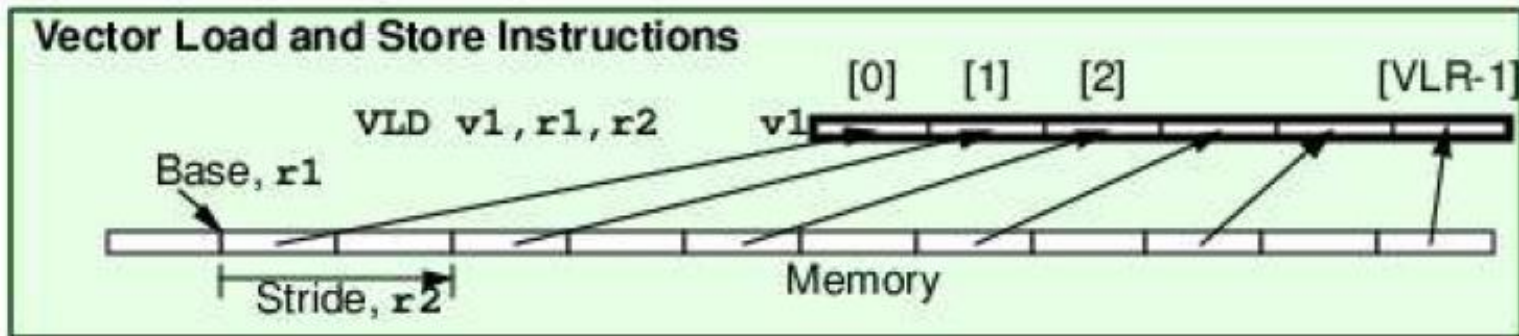
Endereços são encaminhados ao Controlador de Memória e Gerador de Endereços, onde os operandos são buscados

A Unidade Aritmética Pipeline executa as instruções vetoriais

Arquiteturas SIMD

Processadores Vetoriais

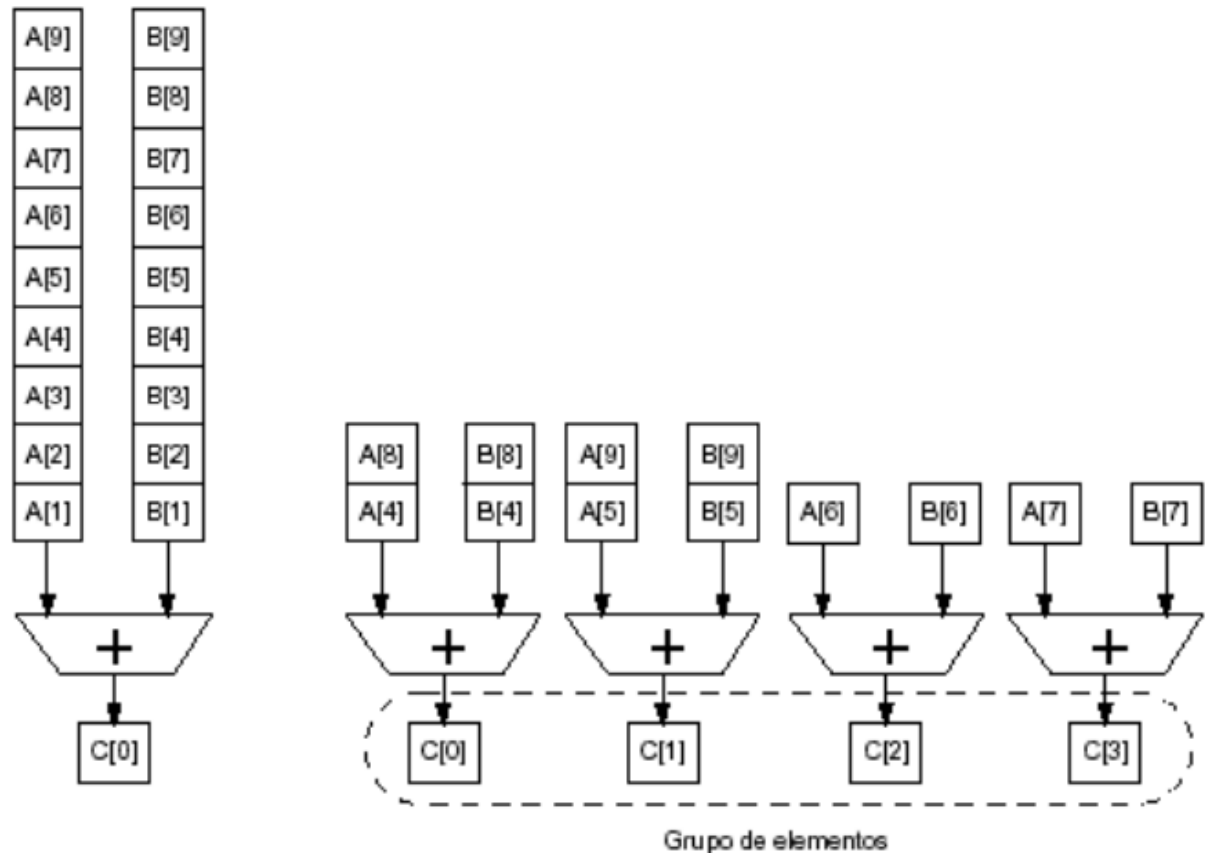
Exemplo de uma instrução de carregamento de um vetor



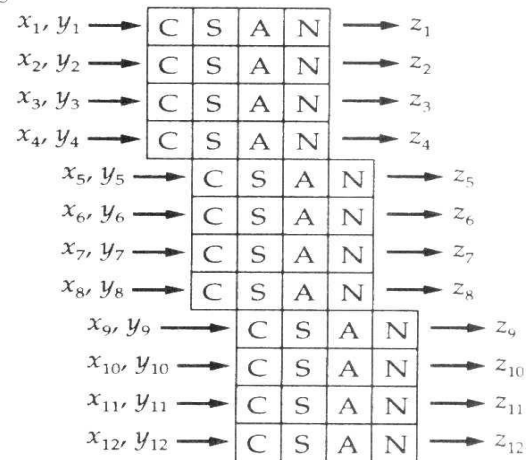
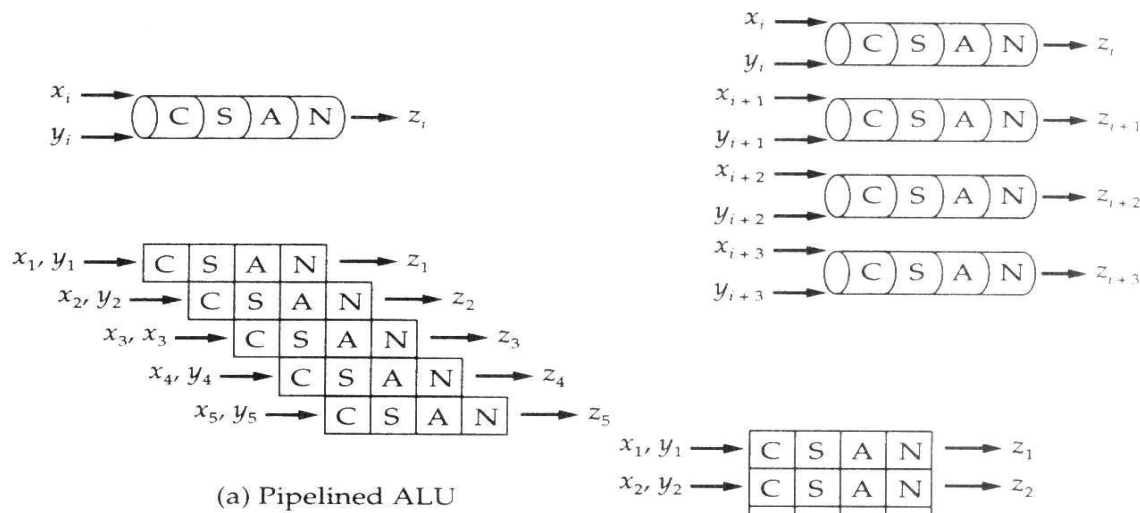
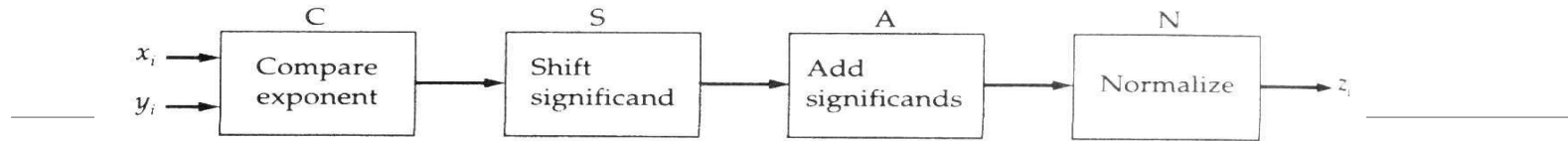
Arquiteturas SIMD

Processadores Vetoriais

Para melhorar ainda mais o desempenho, pode-se aumentar o número de ULAs



Processamento vetorial – *pipelines*



Arquiteturas SIMD

Processadores Vetoriais

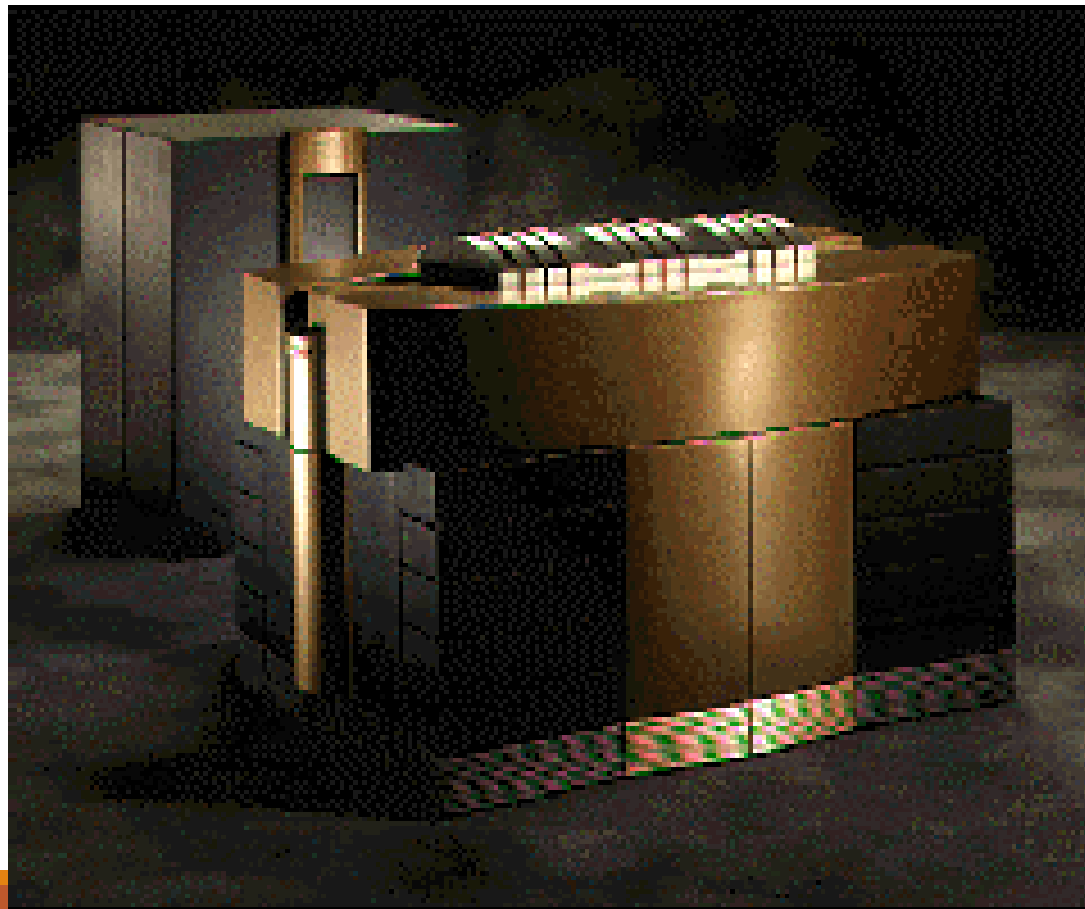
Problemas para a vetorização

- Quando o tamanho do vetor é maior do que o dos registradores vetoriais, o compilador deve quebrar em várias operações vetoriais
- Stride
 - Quantidade de dados que devem ser “saltadas” na leitura/escrita
 - Utilizado principalmente quando se trabalha com matrizes, pois estas não são armazenadas sequencialmente na memória
 - Normalmente as instruções de *load* e *store* possuem um valor para se especificar o stride

Arquiteturas SIMD

Processadores Vetoriais

Supercomputador Vetorial Cray T90



Instruções Vetoriais em Microprocessadores

MMX (Pentium | |)

- Registradores Vetoriais de 64 bits

3DNow (AMD)

SSE (Streaming SIMD Extension) (Pentium III)

- 8 Registradores Vetoriais de 128 bits e novo conjunto de Instruções (+70)

SSE2 (Pentium IV)

- Registradores Vetoriais de 128 bits com elementos de 64 bits mais novas instruções (+144)

SSE3 (Pentium IV), SS4 (Intel Core), AVX (*Advanced Vector Extensions*)

Instruções Vetoriais em Microprocessadores

Problemas no uso de Instruções Vetoriais em Microprocessadores

- Vetores curtos e poucos registradores
- Modos de endereçamento de memória simples (espaçamento único)
- Falta de suporte em linguagem de programação para operar sobre dados estreitos(elementos)
- Consequentemente, Instruções SIMD geralmente são encontradas apenas em bibliotecas de rotinas gráficas de baixo nível

Arquiteturas SIMD

Arquiteturas Matriciais

Processadores executam sincronizadamente a mesma instrução sobre dados diferentes.

Utiliza vários processadores especiais muito mais simples, organizados em geral de forma matricial.

Muito eficiente em aplicações onde cada processador pode ser associado a uma sub-matriz independente de dados (processamento de imagens, algoritmos matemáticos, etc.)

Não incluem processamento escalar

Arquiteturas SIMD

Processadores Matriciais

Conectividade

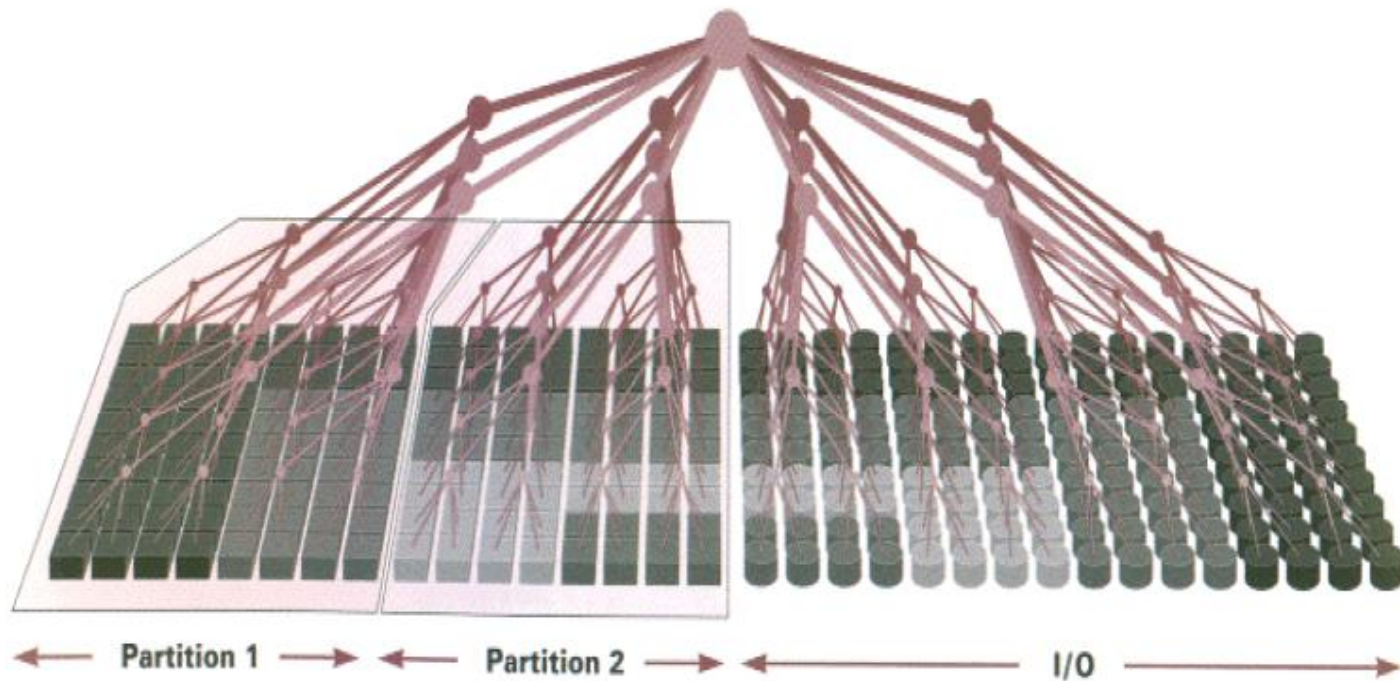
- Vizinhança
- Árvore
- Pirâmide
- Hipercubo

Arquiteturas SIMD

Processadores Matriciais

CM-5

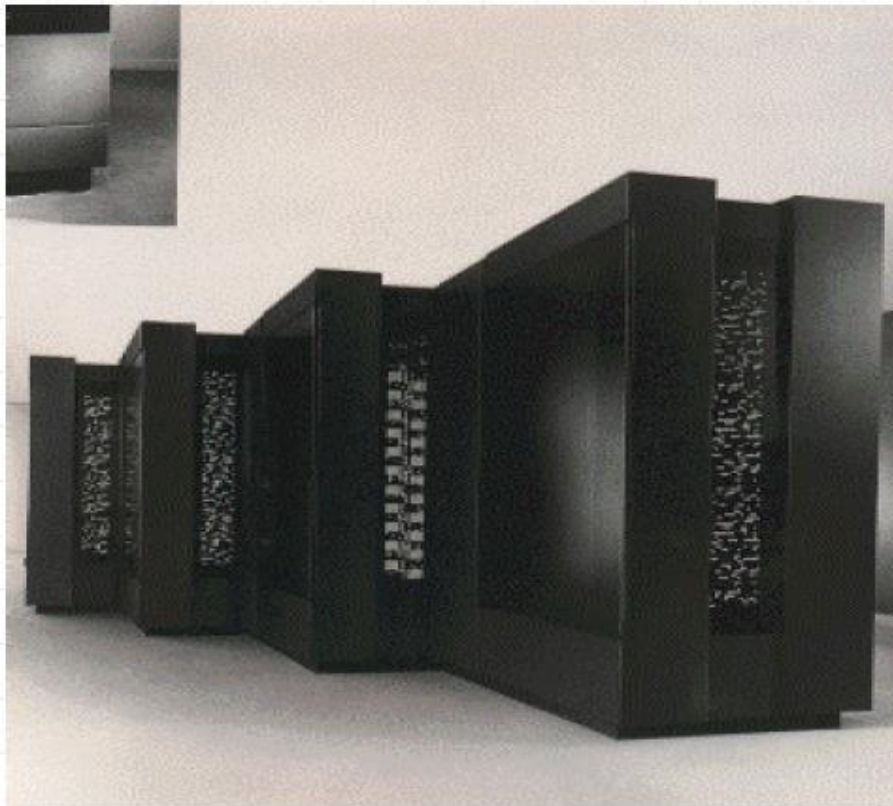
- Conexão em árvore



Arquiteturas SIMD

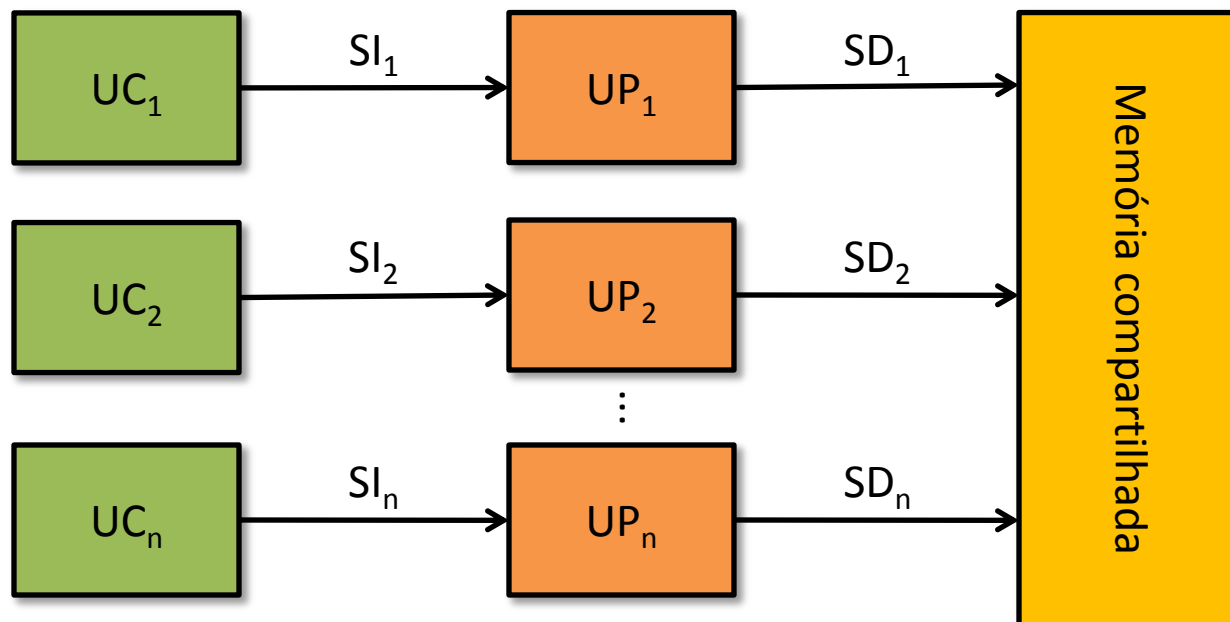
Processadores Matriciais

CM-5 (Connection Machine 5)



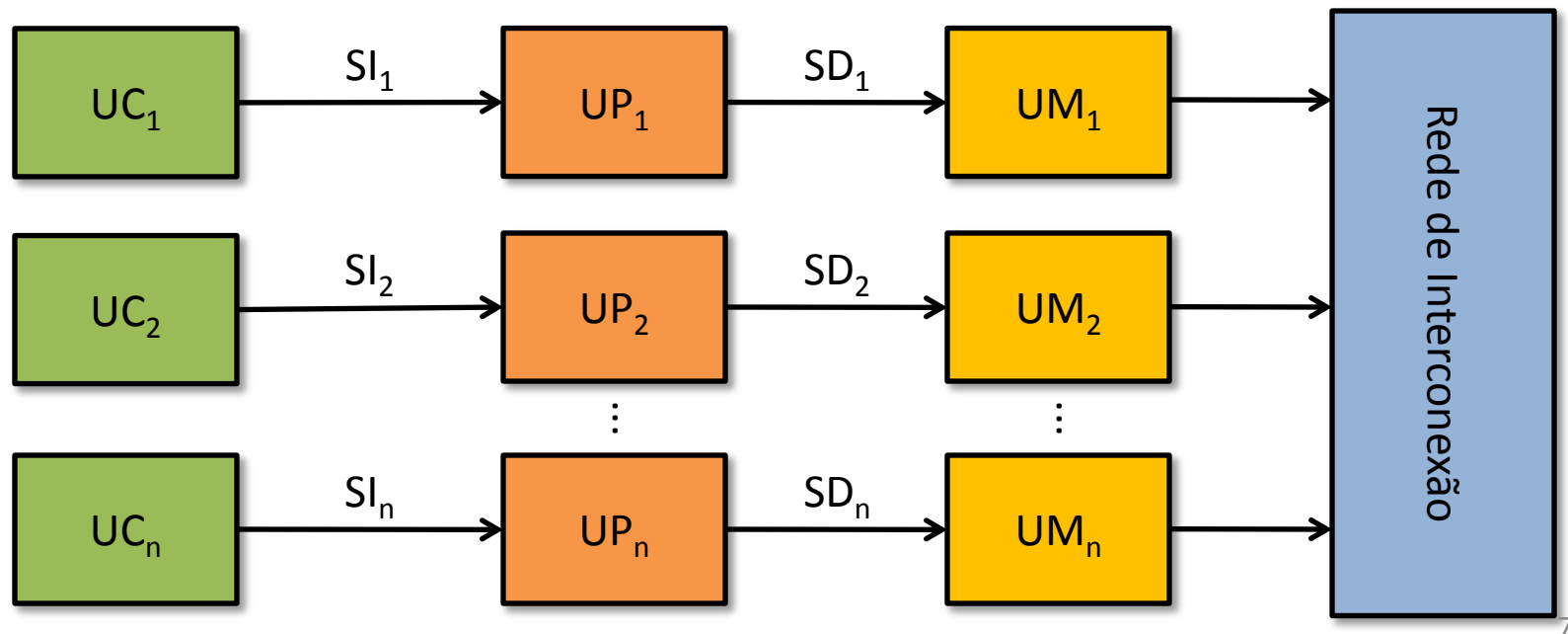
Organizações de Múltiplos Processadores

- MIMD (Múltiplas Instruções, Múltiplos Dados)
 - Há múltiplas UC, cada qual alimentando sua própria UC com uma SI diferente



Organizações de Múltiplos Processadores

- MIMD (Múltiplas Instruções, Múltiplos Dados)
 - Há múltiplas UC, cada qual alimentando sua própria UC com uma SI diferente



PROCESSAMENTO PARALELO

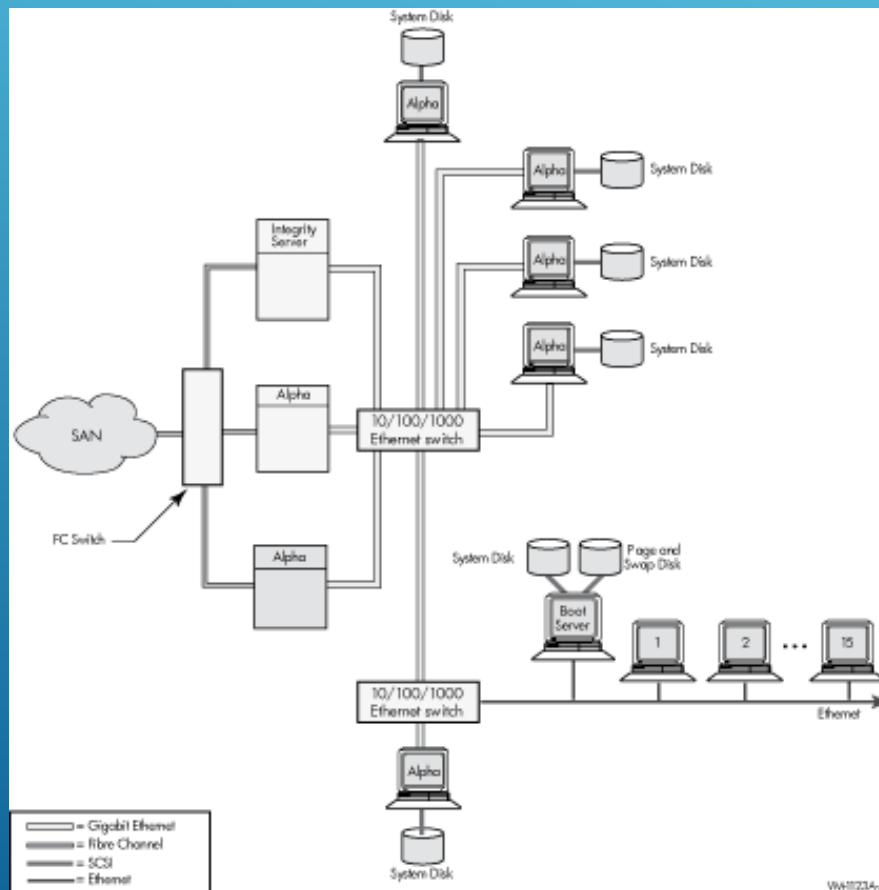
Cluster de Workstations:

Um *Cluster* de Computadores, também conhecido como *NOW* (*Network of Workstations*) ou *COW* (*Cluster of Workstations*), pode ser visto, em sua forma mais simples, como um conjunto composto por dois ou mais computadores (monoprocessáveis ou multiprocessáveis) que trabalham em harmonia buscando fornecer uma solução para um problema, geral ou específico. É uma tecnologia capaz de substituir supercomputadores em certos tipos de aplicações (como simulações, cálculos científicos entre outras) que necessitam de alto desempenho, porém com um custo substancialmente menor, utilizando processadores/computadores com menor poder computacional individual e preço mais baixo.

Os três tipos mais comuns de *clusters* encontrados atualmente são os de Alta Disponibilidade (*High Availability – HA*), de Alto Desempenho de Computação (*High Performance Computing – HPC*) e de Balanceamento de Carga (*Horizontal Scaling – HS*).

PROCESSAMENTO PARALELO

Cluster de Workstations:



PROCESSAMENTO PARALELO

Exemplo de Cluster de Workstations:

O SETI@home é um experimento científico que utiliza a capacidade de processamento ociosa de computadores conectados à Internet de voluntários para a busca de inteligência extraterrestre através da análise de sinais do espaço obtidos de um rádio-telescópio.

SETI (Search for Extraterrestrial Intelligence), ou Busca por Inteligência Extraterrestre, é uma área da ciência dedicada à busca de vida inteligente fora do nosso planeta. A abordagem do SETI@home é conhecida como "rádio SETI", ou seja, busca de sinais inteligentes através de sinais de rádio.

PROCESSAMENTO PARALELO

Exemplo de Cluster de Workstations:

Os sinais analisados pelo SETI@home são obtidos principalmente pelo rádio-telescópio de Arecibo, em Porto Rico, e que é o maior rádio-telescópio fixo do mundo, com uma antena de 305 metros de diâmetro. A análise dos sinais de rádio exige uma capacidade de processamento muito grande, assim em 1995 surgiu a idéia de se dividir os sinais captados em pequenos "pacotes" e distribuí-los para processamento por computadores de voluntários ao redor do mundo.



PROCESSAMENTO PARALELO

Exemplo de Cluster de Workstations:

Segundo SETI@home:

BOINC (Berkeley Open Infrastructure for Network Computing), é uma plataforma aberta para projetos de processamento distribuído, como o SETI@home.

O software cliente do BOINC é altamente customizável e permite que você escolha quais projetos vai ajudar e defina quanto do processamento do seu computador será destinado a cada projeto. Além de ajudar a buscar por vida inteligente com o SETI@home, você pode ainda ajudar a vários outros experimentos científicos, por exemplo: busca pela cura de doenças como o câncer e a AIDS, estudos climáticos e de aquecimento global, estudos astronômicos, etc.

O funcionamento do BOINC é basicamente o seguinte:

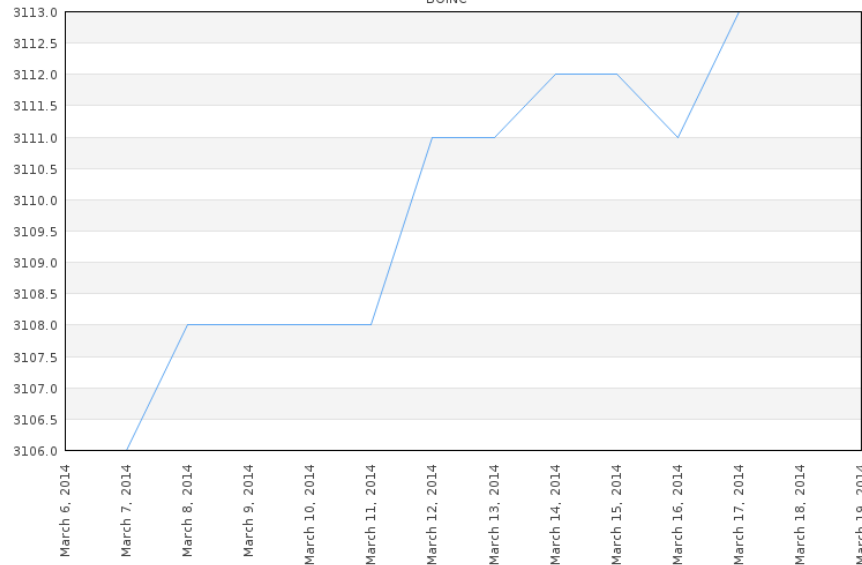
1. após o download e instalação do BOINC, você define de quais projetos vai participar. Você pode definir preferências específicas para cada projeto, por exemplo para limitar o uso de processamento, memória e espaço em disco;
2. o BOINC conecta aos servidores dos projetos selecionados e faz o download das "aplicações" dos projetos, que são basicamente os programas para processamento das informações daquele projeto em específico;
3. o BOINC conecta aos servidores dos projetos e faz o download das "tarefas", que são basicamente os dados a serem processados;
4. Após o processamento das tarefas, o BOINC retorna os resultados para os servidores dos projetos.

O BOINC faz automaticamente as atualizações das aplicações dos projetos dos quais você participa e o download de novas tarefas para processamento. Assim, após a configuração básica, você não precisará fazer nada, e terá um software que funciona como um screensaver (protetor de tela) e que ajuda a ciência.

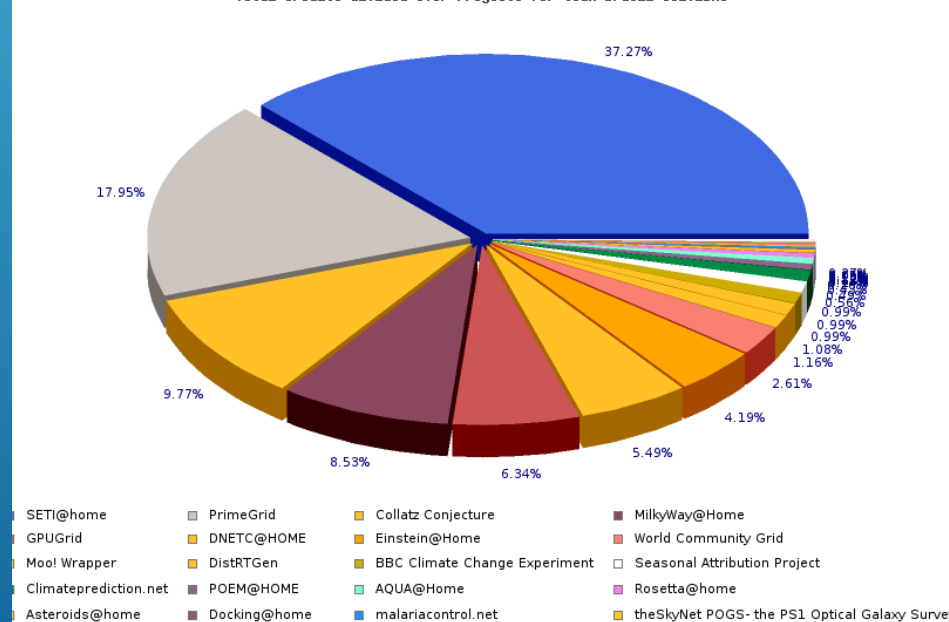
PROCESSAMENTO PARALELO

Exemplo de Cluster de Workstations:

Members in team Brasil [SETIBR]
-BOINC-

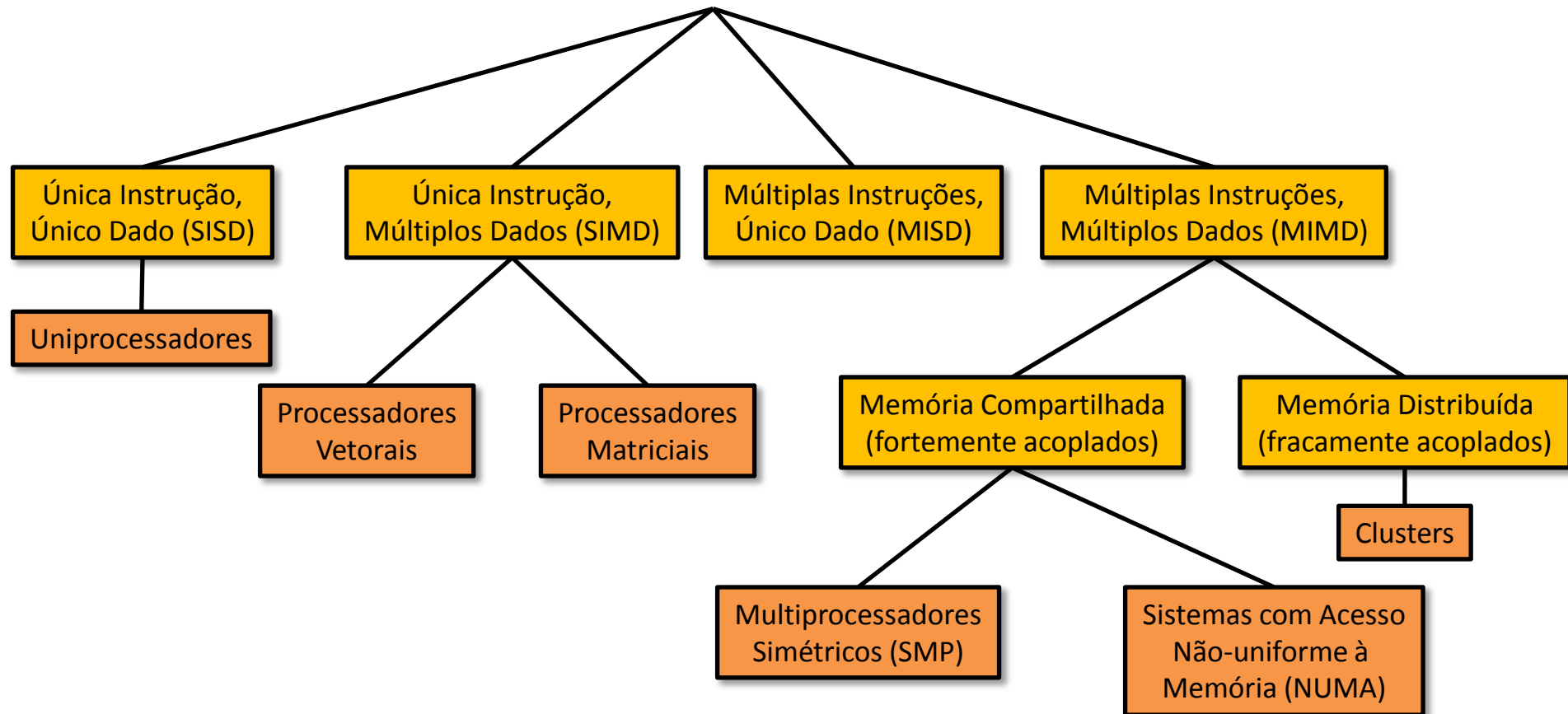


Total credits divided over Projects for team Brasil [SETIBR]

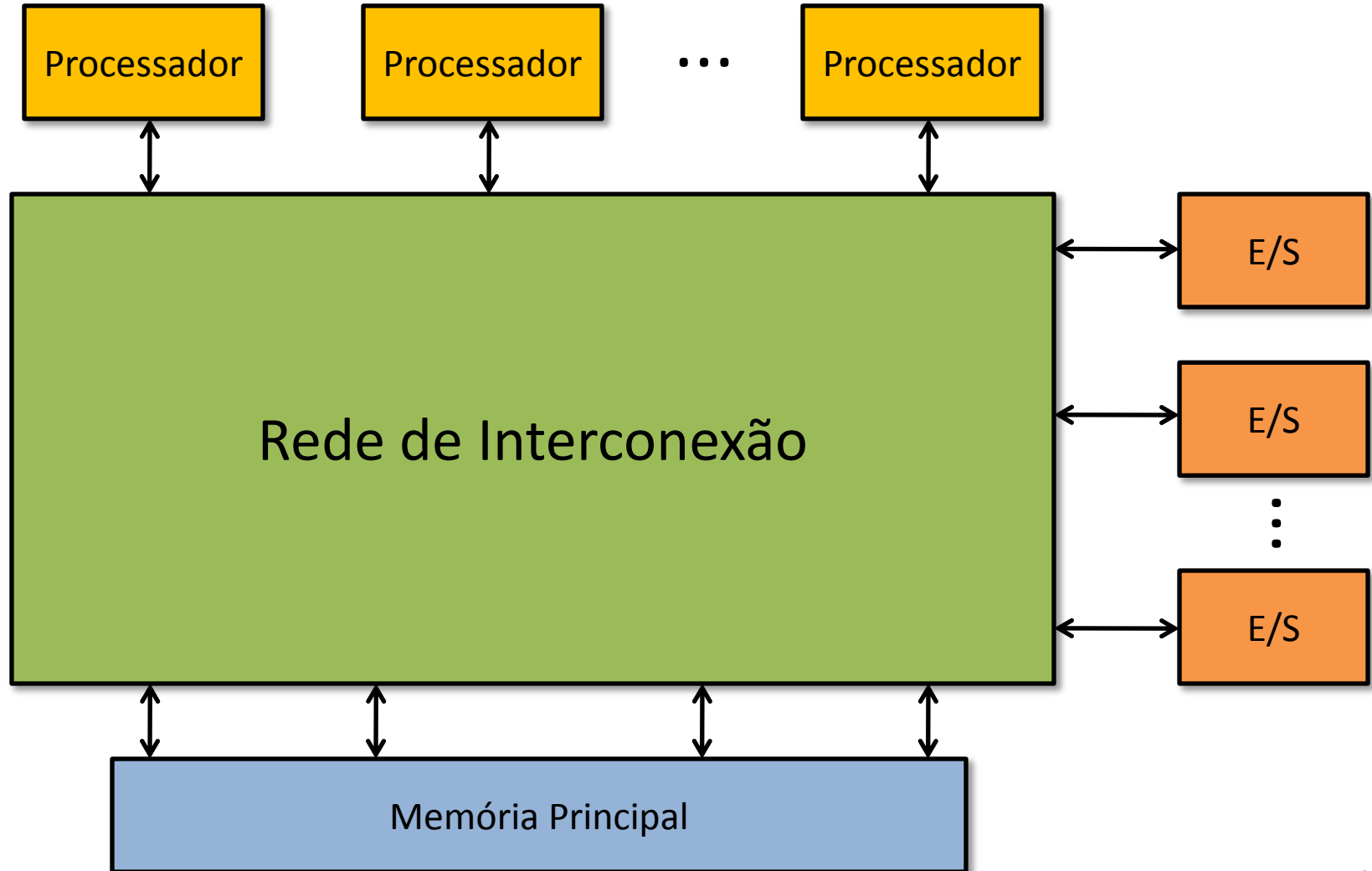


Organizações de Múltiplos Processadores

Classificação



Multiprocessadores Simétricos



Multiprocessadores Simétricos

- Vantagens **potenciais** sobre uma arquitetura uniprocessador
 1. Desempenho
 2. Disponibilidade
 3. Crescimento incremental
 4. Personalização do fornecimento

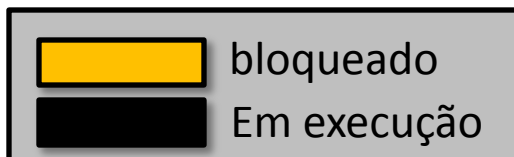
Multiprocessadores Simétricos



(a) Tempo compartilhado (multiprogramação)



(b) Tempo compartilhado e sobreposição (multiprocessamento)



Multiprocessadores Simétricos

- Classificação da organização
 1. Tempo compartilhado ou barramento comum
 2. Memória com múltiplas portas
 3. Unidade de controle central