

# **William Stallings**

# **Computer Organization**

# **and Architecture**

---

## **Chapter 13**

## **Processamento Paralelo**

# Introdução

---

- Tradicionalmente, o computador tem sido visto como uma máquina seqüencial
- Esta visão nunca foi completamente verdadeira
  - No nível das  $\mu O$ , 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)

# Organização de Múltiplos Processadores

---

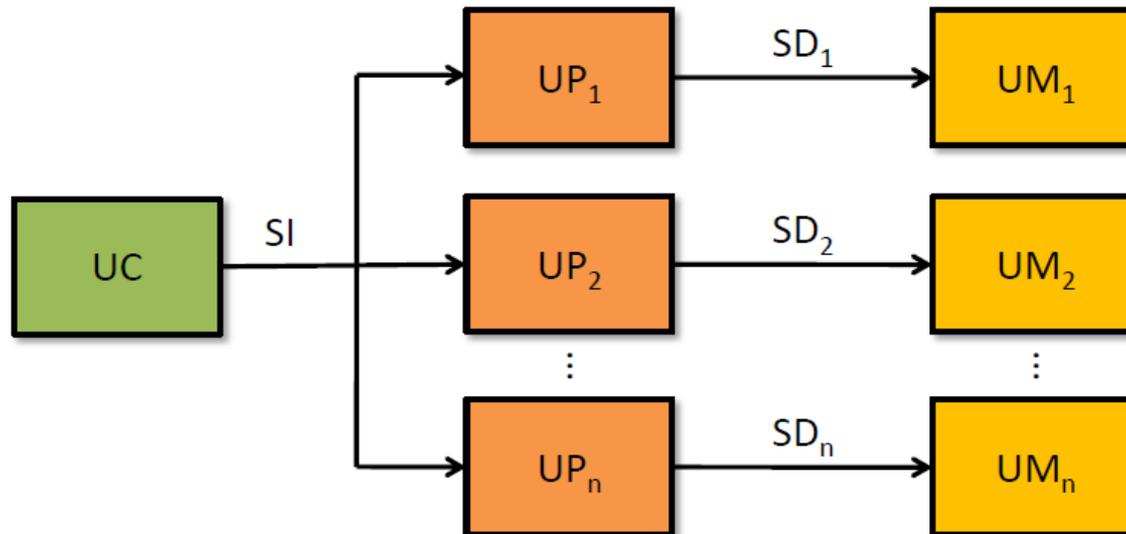
- SISD (Única Instrução, Único Dado)
  - Unidade de Controle (UC) fornece uma seqüência de instruções para a Unidade de Processamento (UP)
  - UP opera sobre uma única Seqüência de Dados (SD) de uma única Unidade de Memória (UM)
  - Exemplo: sistema uniprocessado



# Organização 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



# Arquitetura 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

## 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.

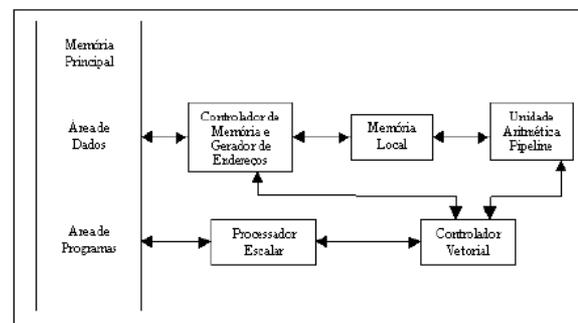
# Arquitetura SIMD – Processadores Vetoriais

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

## 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



# Arquitetura SIMD – Processadores Vetoriais

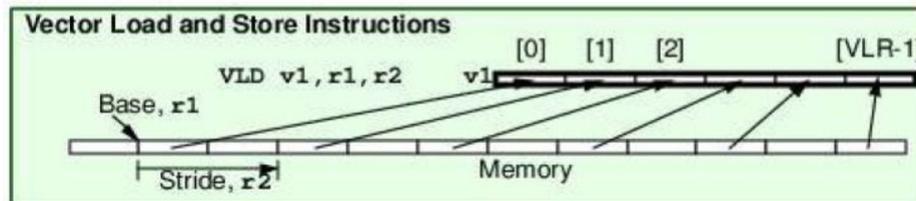
Processador Escalar:

- 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

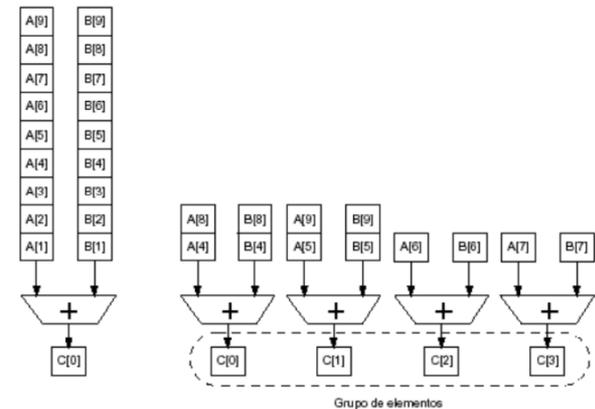
Endereços são encaminhados ao Controlador de Memória e Gerador de Endereços, onde os operandos são buscados

A Unidade Aritmética Pipeline executa as instruções vetoriais

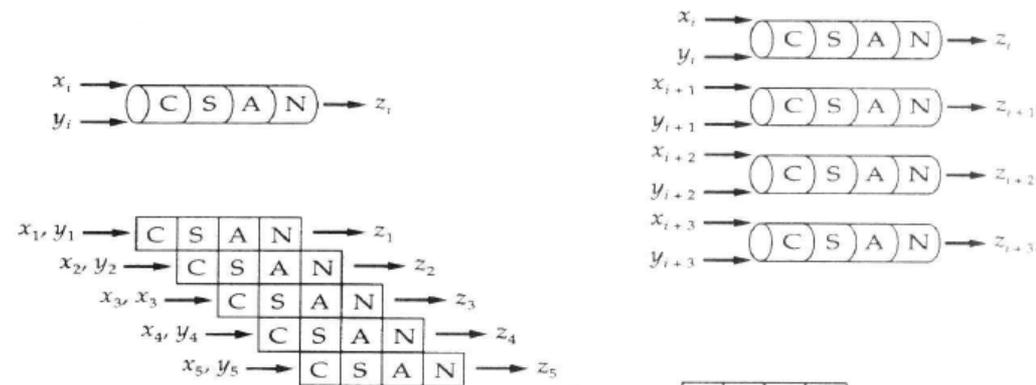
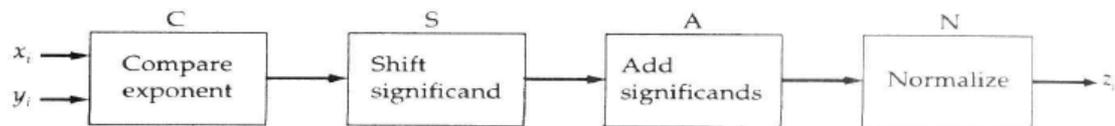
Exemplo de uma instrução de carregamento de um vetor



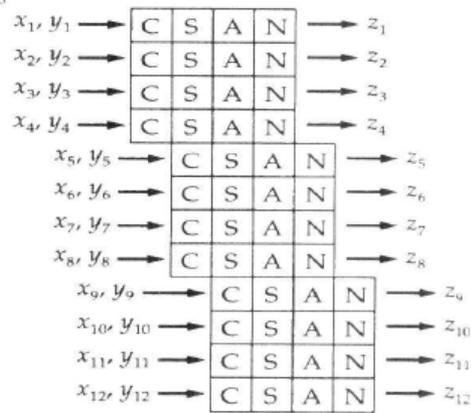
Para melhorar ainda mais o desempenho, pode-se aumentar o número de ULAs



# Arquitetura SIMD – Processadores Vetoriais - Pipelining



(a) Pipelined ALU



(b) Four parallel ALUs

# Arquitetura SIMD – Processadores Vetoriais

---

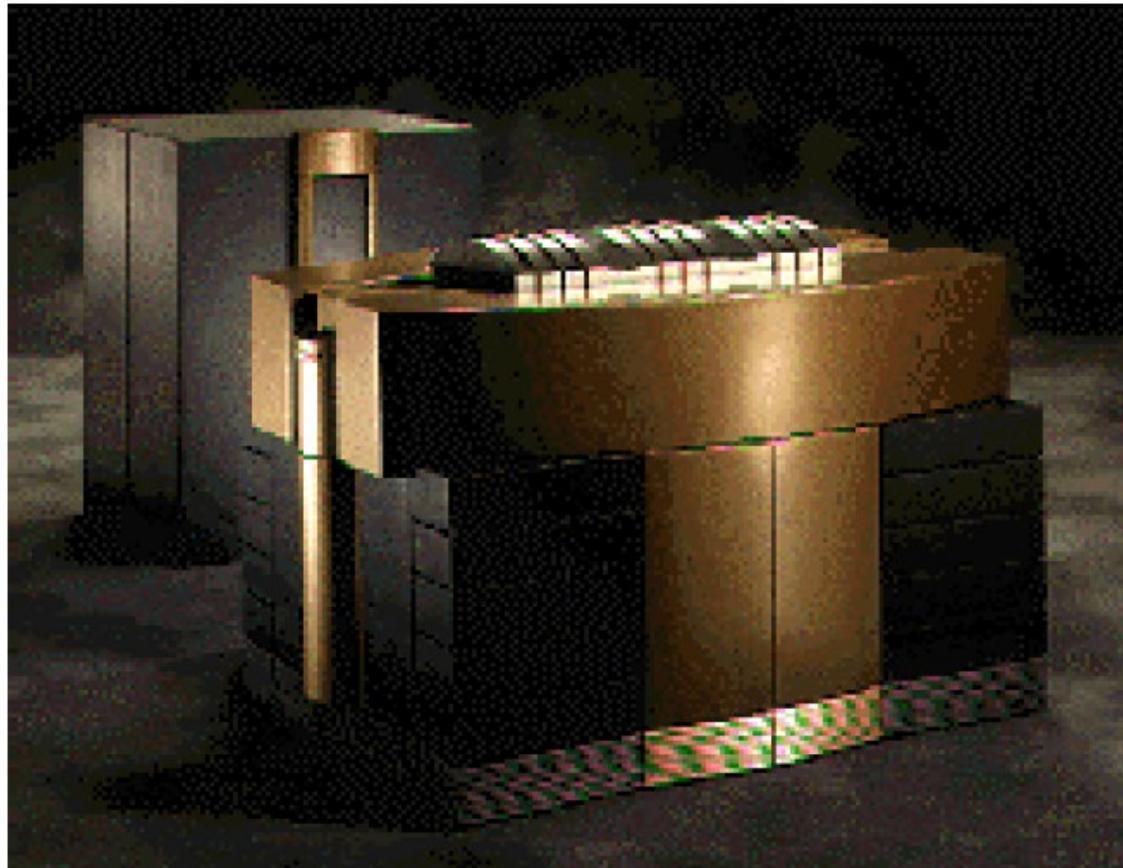
## Problemas para a vetorização

- 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

# Arquitetura SIMD – Processadores Vetoriais

---

Supercomputador Vetorial Cray T90



# Arquitetura SIMD – 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*)

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)
- Conseqüentemente, Instruções SIMD geralmente são encontradas apenas em bibliotecas de rotinas gráficas de baixo nível

# Arquitetura 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.)

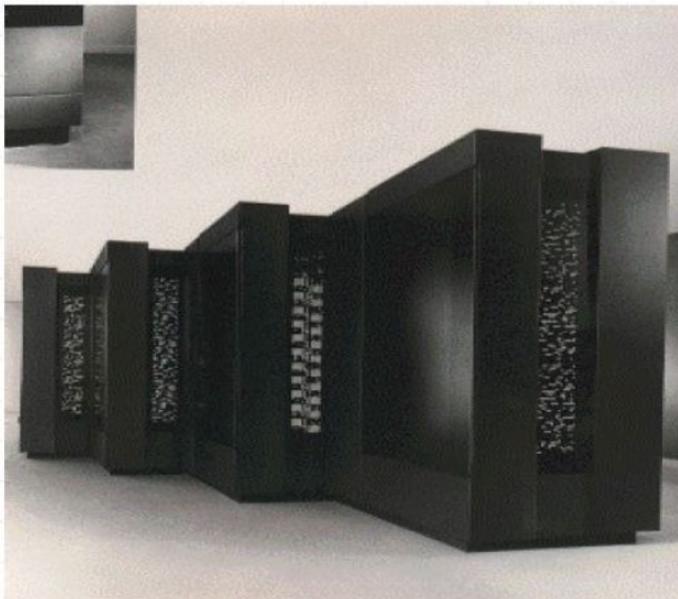
Não incluem processamento escalar

Conectividade

- Vizinhança
- Árvore
- Pirâmide
- Hipercubo

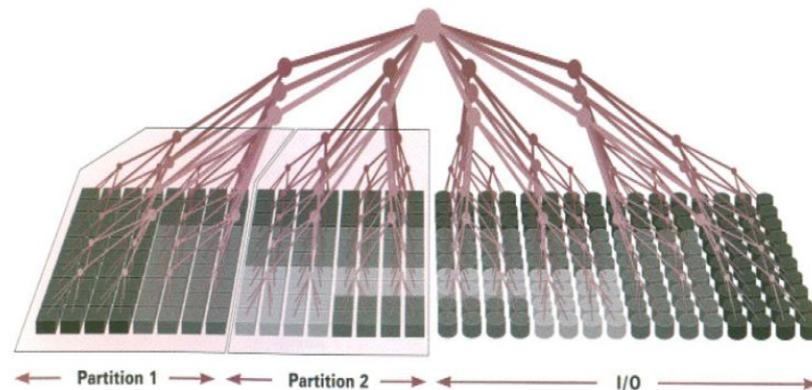
# Arquitetura SIMD – Arquiteturas Matriciais

CM-5 (Connection Machine 5)



CM-5

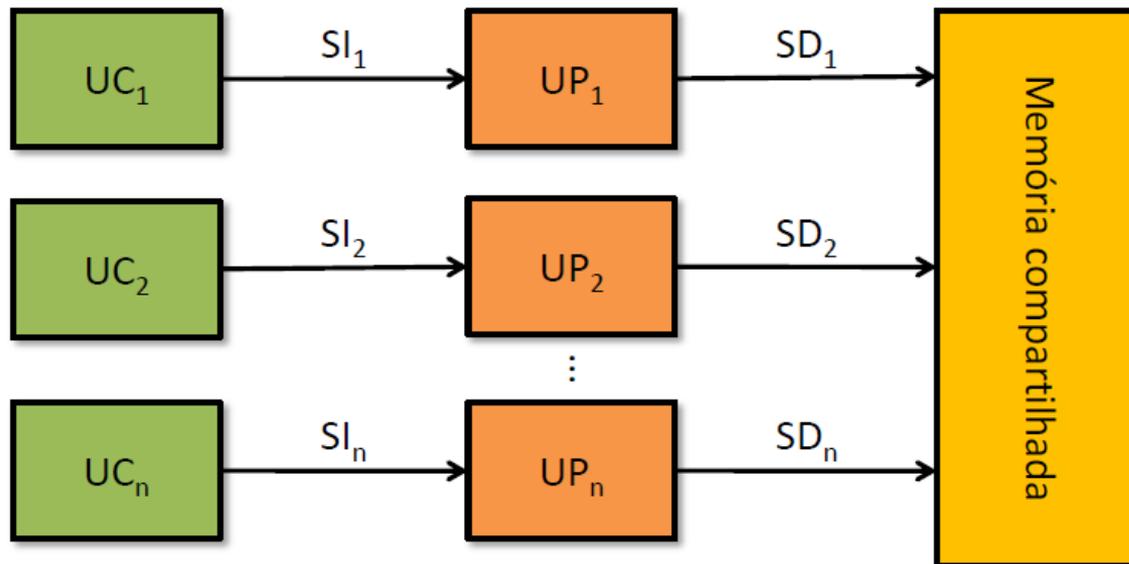
- Conexão em árvore



# Organização de Múltiplos Processadores

---

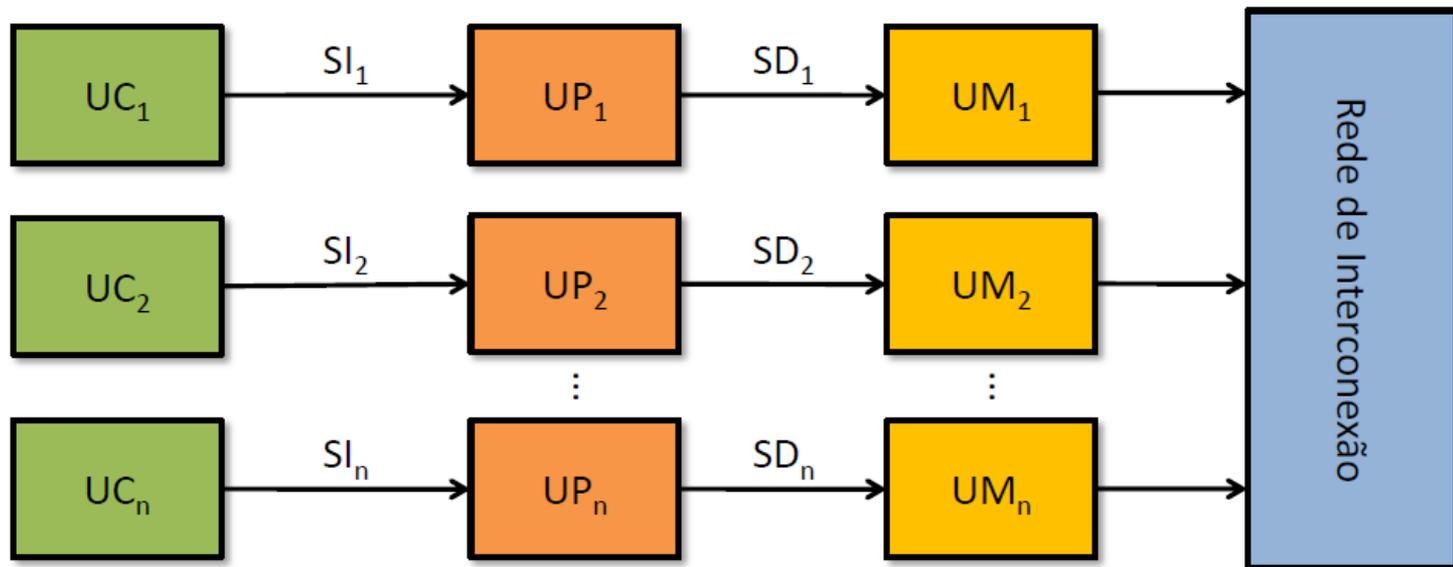
- MIMD (Múltiplas Instruções, Múltiplos Dados)
  - Há múltiplas UC, cada qual alimentando sua própria UC com uma SI diferente



# Organização de Múltiplos Processadores

---

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



# Processamento Paralelo

---

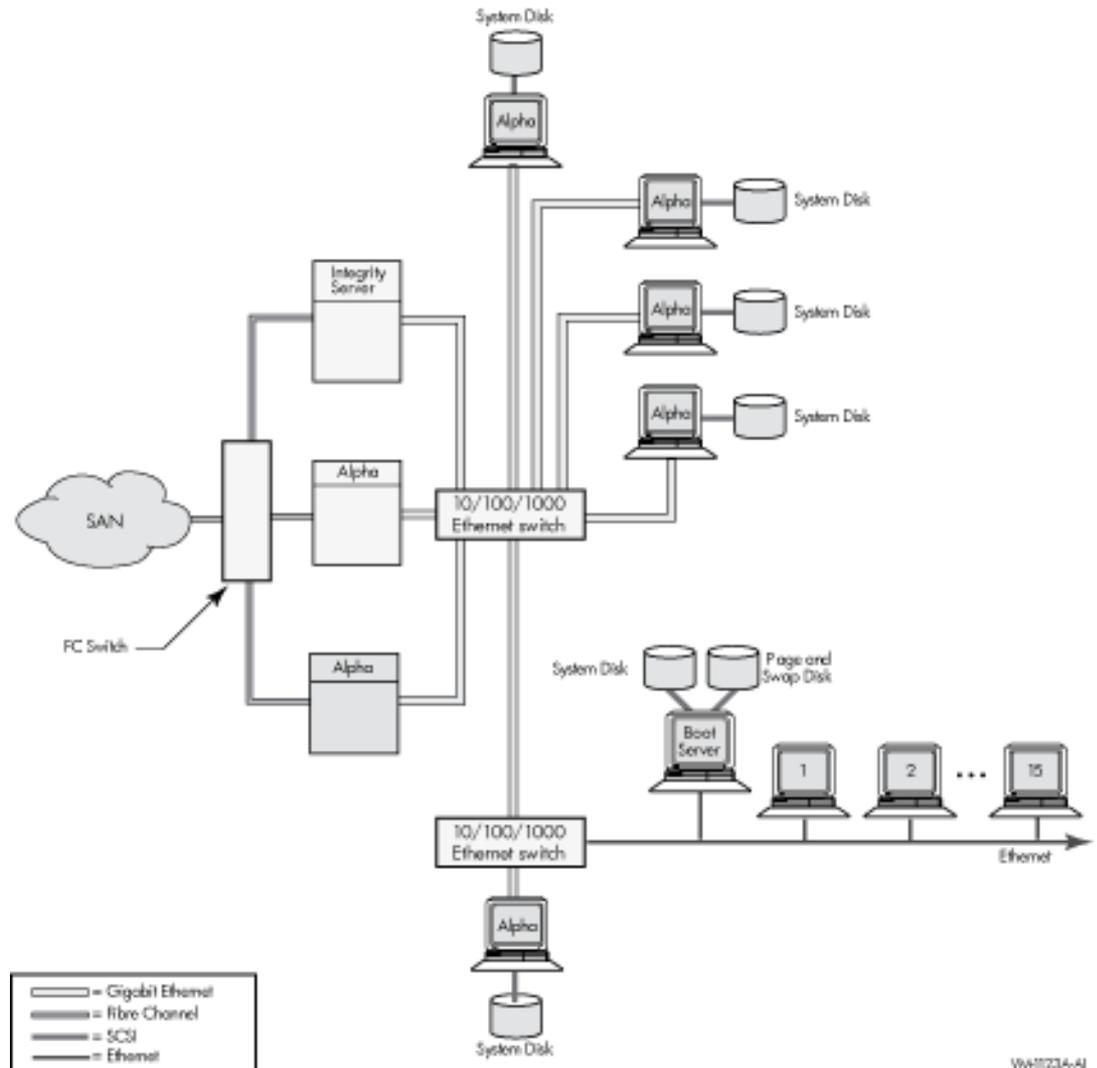
## Cluster de Workstations:

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:



### Introduction

BOINC is a software system for "volunteer computing": it lets people donate time on their home computers and smartphones to science research projects. It has been used by about 50 projects in many areas of science, and has run on millions of computers.

Computing has emerged as a central tool in every area of science. BOINC sought to expand this tool by orders of magnitude, to involve the worldwide public directly in science, and to enable breakthrough science. It succeeded to some extent; BOINC-based computing contributed to over 400 research papers. Hence I think that volunteer computing - and BOINC in particular - is an important chapter in the history of science.

This essay tells the story of BOINC from my perspective. It describes how things unfolded, and offers some theories about why things happened as they did. The story is partly about technology but mostly about people: personalities, aspirations, achievements, and conflicts. If anyone feels slighted, I apologize. If there's an omission or factual error, please let me know.

### The origins of BOINC

#### David Gedye and SETI@home

I taught in the UC Berkeley Computer Science Dept. from 1985 to 1992. David Gedye was a graduate student there, and was a Teaching Assistant in my Operating Systems class. We became running partners, and have been fast friends (ha ha) ever since.

Gedye conceived of the idea of volunteer computing, and he shared it with me in early 1995. Together with Woody Sullivan and Dan Werthimer we formed SETI@home, with the goal of using volunteer computing to analyze radio telescope data, looking for "technosignatures": synthetic signals coming from space.

# 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

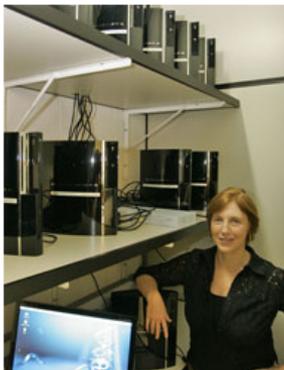
## Exemplo de Cluster de Workstations:

### Unicamp usa PlayStation 3 para realizar pesquisas

Projeto de biologia tem 12 videogames que fazem cálculos científicos. Console da Sony tem melhor custo/benefício que servidores comuns, diz especialista.

RENATO BUENO  
Do G1, em São Paulo

Tamanho da  
letra  
A- A+



AMPLIAR

PS3 ajudam a pesquisadora Monica Pickholz a realizar cálculos de bioinformática (Gustavo Tillo/G1)

O que poderia ser uma LAN house dos sonhos, com o console mais caro da história dos videogames, é, na verdade, um laboratório que faz bilhões de cálculos por segundo para entender melhor a "interação de anestésicos locais com membranas biológicas".

#### Veja galeria de fotos

São 12 PlayStation 3 ligados em rede em uma sala da Unicamp (Universidade Estadual de Campinas), interior de São Paulo. Com seus joysticks (controles) guardados no armário, rodam o sistema operacional Linux, formando um 'cluster' de processamento e ajudando a pesquisadora argentina Monica Pickholz nos cálculos de bioinformática desde junho de 2007.

"São muito mais estáveis que qualquer cluster [aglomerado de PCs] com que já trabalhei", diz Monica. Os videogames funcionam 24 horas por dia, sete dias por semana. Só

pararam uma vez, quando "acabou a força e o gerador não funcionou". Durante todo o tempo, os videogames fazem cálculos para simular dinâmicas de comportamento entre átomos.

### POR DENTRO DO PLAYSTATION 3



#### processamento

Chip: Cell 3.2 GHz, projetado por Sony, IBM e Toshiba. Tem um núcleo e oito processadores. Pode realizar até 2 trilhões de cálculos por segundo.

Memória RAM: 512 MB (256 MB dedicados)

Poder de processamento: 2 TFLOPS\* (o computador mais rápido do mundo atinge 280,6 TFLOPS. Uma calculadora comum consegue 10 FLOPS).

#### armazenamento

Discos rígidos de 20, 60 e 80 GB, dependendo da versão. Conexão via SATA.

Mídia: Blu-ray 25GB (50GB em dupla camada), capaz de rodar jogos e filmes. Também aceita CD e DVD. O disco rígido pode ser preparado para a instalação de sistemas operacionais, como o Linux.

#### conexões de vídeo

Entradas para HDMI, portas analógicas e digitais

Resoluções: 480i, 480p, 720p, 1080i, 1080p

#### entradas

Bluetooth, Wireless, USB, Ethernet

\*FLOPS: "Operações de ponto flutuante por segundo", medida usada para determinar o desempenho de computadores em cálculos científicos.

# Processamento Paralelo

---

## Exemplo de Cluster de Workstations:

### Condor Cluster

Você sabia que já existiu um supercomputador formado por diversas unidades do PS3? É isso mesmo. O **Condor Cluster**, como ficou conhecido, era um conjunto de 1.760 unidades do PlayStation 3 – conectados por cinco quilômetros de fio -, que foi utilizado pela Força Aérea dos EUA em 2010. Esse cluster de consoles da Sony chegou a ocupar a 33ª posição entre os supercomputadores mais poderosos do mundo. A engenhoca envolvia o uso dos consoles em conjunto com 168 unidades de processamento gráfico e 84 servidores coordenadores.



Condor Cluster

A solução também foi considerada uma baita relação custo x benefício, já que o Condor Cluster consumia cerca de 10% da capacidade elétrica quando comparado a um supercomputador nos moldes tradicionais.

# Processamento Paralelo

---

## Exemplo de Cluster de Workstations: COVID19 - High Performance Computing



From left: Quartz, Lassen, and Corona supercomputers, three of the systems being deployed in the fight against COVID-19 by LLNL and HPC consortium researchers.

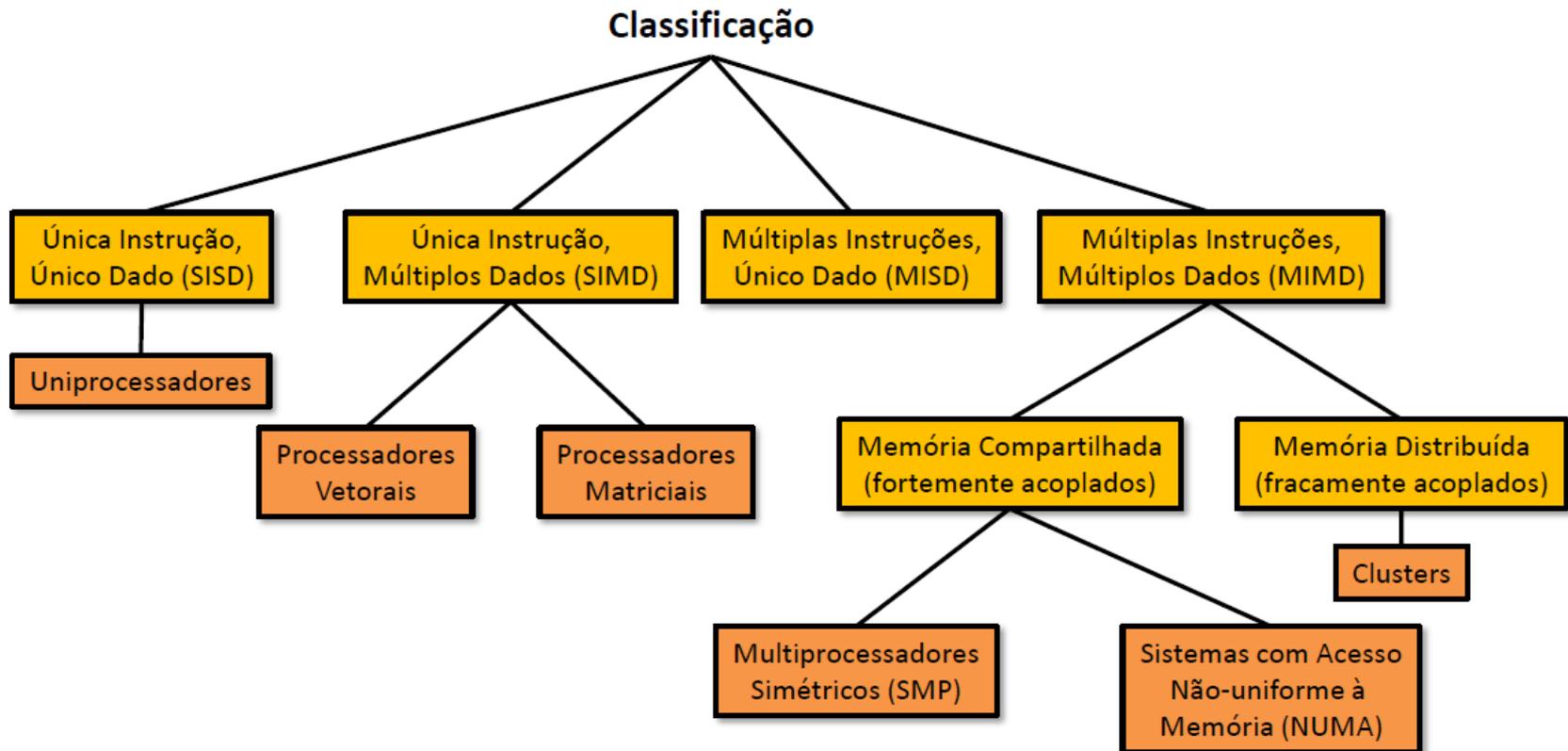
In support of Livermore's research, the Laboratory operates a major computing center, Livermore Computing (LC), one of the most prominent and consistently successful computer centers in the world. Among its many computational assets, LC is home to some of the most powerful computers, several of which are capable of petascale computing ( $10^{15}$  floating point operations per second). These computers use tens of thousands of cores (central and graphics processing units) running at the same time—known as parallel processing.

The Laboratory is participating in a [White House initiative](#) to provide COVID-19 researchers worldwide with access to the world's most powerful high performance computing (HPC) resources that can significantly advance the pace of scientific discovery in the fight to stop the virus. For example, LLNL's Corona supercomputing cluster, which will be used by COVID-19 researchers, is [undergoing upgrades](#) to nearly double the system's peak compute power. (See news links below for information about additional HPC systems dedicated to COVID-19 research.)

Researchers are invited to submit COVID-19 related research proposals to the consortium via the [online portal](#), which will then be reviewed and matched with computing resources from one of the partner institutions. An expert panel of top scientists and computing researchers will coordinate with proposers to quickly assess the public health benefit of the work and coordinate the allocation of the group's powerful computing assets.

# Organização de Múltiplos Processadores

---



# Máquina com Acesso Não Uniforme à Memória

---

Em uma máquina NUMA, a memória é fisicamente composta por vários bancos de memória, podendo estar, cada um, vinculados a um processador. Nesse caso, quando o processador acessa a memória que está vinculada a si diz-se que houve um acesso local.

Se o acesso for à memória de outro processo, ocorre um acesso remoto. Os acessos remotos são mais lentos que os acessos locais já que é necessário passar pela malha de interconexão.

# Multiprocessadores Simétricos (Multicore)

---

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

# Multiprocessadores Simétricos (Multicore)

---



(a) Tempo compartilhado (multiprogramação)



(b) Tempo compartilhado e sobreposição (multiprocessamento)



# Multiprocessadores Simétricos (Multicore)

---

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

# Clusters de Computadores - Supercomputadores

TOP 500  
Junho / 24

Rank	System	Cores	Rmax (PFlop/s)	Rpeak (PFlop/s)	Power (kW)
1	<b>Frontier</b> - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	8,699,904	1,206.00	1,714.81	22,786
2	<b>Aurora</b> - HPE Cray EX - Intel Exascale Compute Blade, Xeon CPU Max 9470 52C 2.4GHz, Intel Data Center GPU Max, Slingshot-11, Intel DOE/SC/Argonne National Laboratory United States	9,264,128	1,012.00	1,980.01	38,698
3	<b>Eagle</b> - Microsoft NDv5, Xeon Platinum 8480C 48C 2GHz, NVIDIA H100, NVIDIA Infiniband NDR, Microsoft Azure Microsoft Azure United States	2,073,600	561.20	846.84	
4	<b>Supercomputer Fugaku</b> - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,630,848	442.01	537.21	29,899
5	<b>LUMI</b> - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE EuroHPC/CSC Finland	2,752,704	379.70	531.51	7,107

# Clusters de Computadores - Supercomputadores

## TOP 500 Junho / 24 Brasil

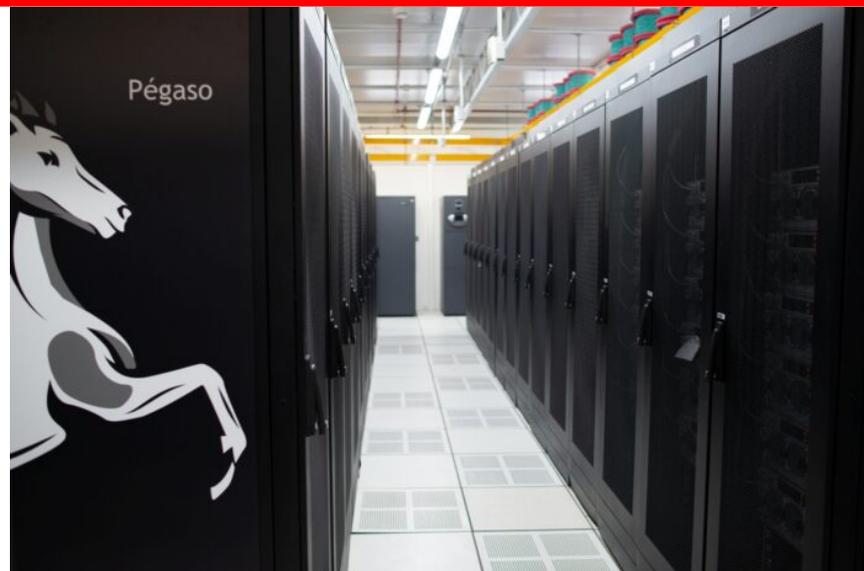
54	<b>Frontier TDS</b> - HPE Cray EX235a, AMD Optimized 3rd Generation EPYC 64C 2GHz, AMD Instinct MI250X, Slingshot-11, HPE DOE/SC/Oak Ridge National Laboratory United States	120,832	19.20	23.11	309
55	<b>Helios GPU</b> - HPE Cray EX254n, NVIDIA Grace 72C 3.1GHz, NVIDIA GH200 Superchip, Slingshot-11, HPE Cyfronet Poland	89,760	19.14	30.44	317
56	<b>Pégaso</b> - Supermicro A+ Server 4124G0-NART+, AMD EPYC 7513 32C 2.6GHz, NVIDIA A100, Infiniband HDR, EVIDEN Petróleo Brasileiro S.A Brazil	233,856	19.07	42.00	1,033

Além de ocupar o posto de supercomputador mais potente da América Latina no TOP500, o Pégaso é líder da região no Green500, ranking de eficiência energética.

Quando plenamente em operação, o supercomputador terá consumo máximo de 1,5 MW. A Petrobras dá a entender que essa é uma capacidade bastante otimizada para uma máquina de tamanho porte. Paulo Palaia, diretor de Transformação Digital e Inovação da empresa, comenta:

Na elaboração do projeto deste HPC, foram feitas várias escolhas e exigências técnicas de maneira a reduzir o consumo, como a escolha de aceleradores matemáticos (GPUs), os mais ecoeficientes que existem, e a exigência de fontes de alta eficiência.

A companhia destaca ainda que o Pégaso foi instalado em um centro de dados que só usa energia limpa, ainda que esse fator não contribua para a pontuação no Green500.



# Clusters de Computadores - Supercomputadores

## Brasil - SDumont

- 504 nós de computação B710 (*thin node*), onde cada nó possui 2 CPU Intel Xeon E5-2695v2 Ivy Bridge (12c @2,4GHz) e 64Gb de memória RAM
- 198 nós de computação B715 (*thin node*) com GPUs K40, onde cada nó possui 2 x CPU Intel Xeon E5-2695v2 Ivy Bridge (12c @2,4GHz) e 64Gb de memória RAM
- 54 nós de computação B715 (*thin node*) com XEON PHI, onde cada nó possui 2 x CPU Intel Xeon E5-2695v2 Ivy Bridge (12c @2,4GHz) e 64Gb de memória RAM
- 1 nó de computação MESCA 2 com memória compartilhada, 16 x CPU Intel Xeon Ivy Bridge (15c @2,4GHz) e 6TB de memória RAM
- 246 nós computacionais (CPU), cada um com 2x Intel Xeon Cascade Lake Gold 6252 e 384Gb de memória RAM
- 36 nós computacionais (CPU), cada um com 2x Intel Xeon Cascade Lake Gold 6252 e 768 Gb de memória RAM
- 94 nós computacionais (GPU), cada um com 2x Intel Xeon Cascade Lake Gold 6252, 4x NVIDIA Volta V100 GPU e 384Gb de memória RAM
- 1 nó para Inteligência Artificial com 2x Intel Xeon Skylake Gold 6148 (20c @2,4GHz), 8x NVIDIA Tesla V100-16GB com NVLink e 384Gb de memória RAM

