

## EGC9001-09

### T.E.E.C. Ambientes Computacionais Distribuídos para Engenharia do Conhecimento

*Prof. Mario Dantas*

#### *Objetivo Geral:*

*Prover um conhecimento teórico das arquiteturas computacionais **não convencionais** em termos de agregação processadores, memória e sistema de interconexão de rede*

*É objetivo da disciplina, também, prover um embasamento necessário para o entendimento de sistemas distribuídos modernos, como os clusters, grids e clouds, em termos de arquitetura computacionais.*

#### *Objetivos Específicos:*

- a) *Compreender a taxonomia de arquiteturas SISD, SIMD, MISD e MIMD, através de seus componentes e funções de processamento e armazenamento;*
- b) *Detalhar ambientes computacionais modernos em termos de seus sistemas de interconexão e formas de acesso a memória uniforme e não uniforme. Exemplificar de forma mais detalhada a arquitetura de multicomputadores, multiprocessadores, máquinas com características SMP, Numa, ccNuma e MPP;*
- c) *Contextualizar as arquiteturas de computadores modernos em ambientes distribuídos tipo clusters, grids e clouds.*

#### *Ementa*

- *Arquitetura de computadores modernos;*
- *Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);*
- *Redes de Interconexão;*
- *Multicomputadores;*
- *Multiprocessadores;*
- *Máquinas com Acesso Uniforme à Memória (UMA);*
- *Multiprocessadores simétricos (SMP);*
- *Máquinas com Acesso Não Uniforme à Memória (NUMA);*
- *Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);*
- *Processadores Massivamente Paralelos (MPP);*
- *Sistemas Distribuídos;*
- *Clusters;*
- *Grids;*
- *Clouds.*

#### *Referências Bibliográficas*

- *David Culler et. al, "Parallel Computer Architecture - A Hardware/Software Approach"*
- *Buyya, R, "High Performance Cluster Computing Vol1. - Architectures and Systems"*
- *Dantas, Mario, "Computação Distribuída de Alto Desempenho: Redes, Clusters e Grids Computacionais"*

#### *Ementa*

- *Arquitetura de computadores modernos;*
- *Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);*
- *Redes de Interconexão;*
- *Multicomputadores;*
- *Multiprocessadores;*
- *Máquinas com Acesso Uniforme à Memória (UMA);*
- *Multiprocessadores simétricos (SMP);*
- *Máquinas com Acesso Não Uniforme à Memória (NUMA);*
- *Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);*
- *Processadores Massivamente Paralelos (MPP);*
- *Sistemas Distribuídos;*
- *Clusters;*
- *Grids;*
- *Clouds.*

### Arquitetura dos Computadores Modernos

#### Objetivo:

Os aspectos da evolução tecnológica na área da arquitetura dos computadores, podem prover uma melhor visão aos profissionais envolvidos em projetos de redes, suporte e desenvolvimento de aplicações distribuídas.

### Arquitetura dos Computadores Modernos

Os primeiros computadores eletrônicos:

(a) J. Presper Eckert e John Mauchly construíram o primeiro computador eletrônico na Moore School na Universidade da Pennsylvania durante a II Guerra Mundial.

### Arquitetura dos Computadores Modernos

A máquina chamada de ENIAC - Electronic Numerical Integrator and Calculator- foi somente conhecida do público em 1946.

Esta máquina foi usada pelo US Army para cálculo das tabelas de tiro.

### Arquitetura dos Computadores Modernos

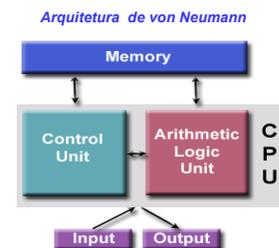
Em 1944, von Neumann foi convidado a participar do projeto. Nesta época o grupo discutia como armazenar os programas para ser processados.

### Arquitetura dos Computadores Modernos

Von Neumann escreveu um artigo sobre o trabalho chamando o computador de EDVA - Electronic Discrete Variable Automatic Computer. Deste fato, surgiu o termo famoso:

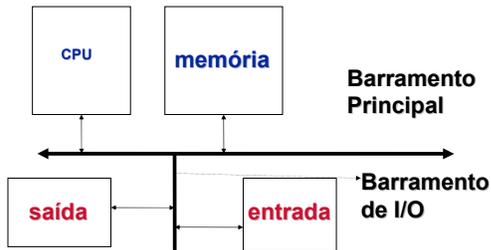
#### Arquitetura de von Neumann

### Arquitetura dos Computadores Modernos



## Arquitetura dos Computadores Modernos

### Arquitetura de von Neumann



## Arquitetura dos Computadores Modernos

Os primeiros Computadores Eletrônicos:

(b) Em 1946, Maurice Wilkes da Universidade de Cambridge visitou a Moore School para assistir aulas sobre o desenvolvimento de computadores eletrônicos. Voltando para Cambridge decidiu criar um projeto semelhante, e criou o EDSAC (Electronic Delay Storage Automatic Calculator).

## Arquitetura dos Computadores Modernos

Os primeiros Computadores Eletrônicos:

O EDSAC foi o primeiro computador eletrônico operacional quanto ao armazenamento de programas.

- (c) Konrad Zuse na Alemanha no final dos anos 30 e princípios dos anos 40 desenvolveu um computador programável.
- (d) Outra iniciativa foi o Colossus desenvolvido pelos ingleses durante a II Guerra.

## Arquitetura dos Computadores Modernos

Eckert and Muchly formaram uma empresa em 1947 para a construção do primeiro computador comercial, o *Binac*.

Depois de problemas financeiros, estes venderam a Remington-Rand que colocou no mercado o UNIVAC I (Universal Automatic Computer). O número de 48 sistemas foram construídos e o custo do primeiro em 1958 foi de US\$ 1.000.000.



UNIVAC1

## Arquitetura dos Computadores Modernos

A IBM estava no mercado de escritórios, mas não antes de 1950 decidiu investir na construção do seu primeiro computador, IBM/701 (1952). Só em 1964, num comunicado surpreendente para a época anunciou o lançamento do System/360. Uma máquina que podia variar em sua configuração e preço.



IBM 360

## Arquitetura dos Computadores Modernos

A Digital por volta de 1965 começa a comercializar o PDP-8, que foi o primeiro *minicomputador* do mercado. Este computador foi uma boa notícia para o mercado de usuários, uma vez que esta máquina podia ser comprada por US\$ 20.000. Somente em 1971 a Intel apresentou o primeiro microprocessador do mercado, o Intel 4004.



PDP1

### Arquitetura dos Computadores Modernos

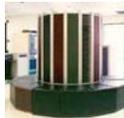
Em 1963 Seymour Cray anunciou o primeiro *supercomputador*, o CDC 6600. Em 1976, Cray já na sua empresa anuncia a máquina mais rápida e cara do mundo, o *Cray-I*.

Em 1996, a SGI (Silicon Graphics) compra a Cray Research o que indica a não existência de mais nenhuma empresa no mercado dedicada *exclusivamente* a construção de *supercomputadores*.

CDC 7600



CRAY1



### Arquitetura dos Computadores Modernos

Em 1977, a computação pessoal é alcançada através do *Apple II* de Steven Jobs e Steve Wozniak.



### Arquitetura dos Computadores Modernos

Devido ao baixo custo, grande volume de armazenamento para a época e a alta confiabilidade estabeleceu-se a indústria dos computadores pessoais. Somente quatro anos depois, 1981, a IBM lança o IBM-PC com o processador Intel e o DOS da Microsoft.



### Arquitetura dos Computadores Modernos

Devido a arquitetura aberta do IBM-PC, logo este microcomputador tornou-se padrão no mercado. A Apple, mesmo com seu Macintosh ficou com um segundo plano muito abaixo do padrão IBM-PC.



### Arquitetura dos Computadores Modernos

(Fonte: Hennessy e Patterson)

Geração	Período	Tecnologia	Produto
I	1950-59	Tubos a vácuo	Computador
II	1960-68	Transistors	Computador baixo custo
III	1969-77	C. Integrados	Minis
IV	1978-	LSI e VLSI	PC e WS

### Arquitetura dos Computadores Modernos

(Fonte: Museu do Computador - Boston)

Ano	Nome	Memória (k)	Preço (US\$)
51	Univac I	48	1.000.000
64	IBM /360	64	1.000.000
65	PDP-8	4	16.000
76	Cray I	32768	4.000.000
81	IBM PC	256	3.000
91	HP9000	16384	7.400
96	Intel Pro	16384	4.400







## Taxonomia de Arquitetura de Computadores

A classificação de *arquiteturas computacionais* mais aceita na comunidade acadêmica e indústria é a conhecida *taxonomia de Flynn*.

## Taxonomia de Arquitetura de Computadores *Taxonomia de Flynn*

O pesquisador *Michael J. Flynn* propôs em 1966, uma classificação dos computadores, baseando-se no fluxo de instruções e no tratamento dos dados.

A *Taxonomia de Flynn* leva em consideração como o fluxo de instruções e como os dados são processados.

	Single Instruction	Multiple Instruction
Single Data	<i>SISD</i>	<i>MISD</i>
Multiple Data	<i>SIMD</i>	<i>MIMD</i>

## Taxonomia de Arquitetura de Computadores *Taxonomia de Flynn*

### *SISD (Single Instruction Single Data)*

Computadores com esta característica são aqueles que executam uma instrução de um programa por vez, ou seja, o modelo tradicional do processador único.

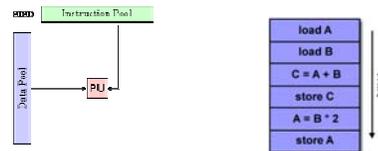
Um exemplo seria seu computador pessoal com um processador *convencional*;

Exemplo:



## Taxonomia de Arquitetura de Computadores *Taxonomia de Flynn*

### *SISD (Single Instruction Single Data)*



## Taxonomia de Arquitetura de Computadores *Taxonomia de Flynn*

### *SIMD (Single Instruction Multiple Data)*

Neste tipo de arquitetura existe, também, a execução de uma única instrução. Todavia, devido à existência de facilidades em hardware para armazenamento, a mesma instrução é processada sob diferentes itens de dados.

## Taxonomia de Arquitetura de Computadores *Taxonomia de Flynn*

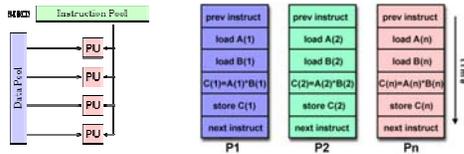
### *SIMD (Single Instruction Multiple Data)*

Quanto as facilidades de hardware para armazenamento, essas normalmente são classificadas como:

- *Processor Arrays*;
- *Vector Pipeline*.

## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

SIMD (Single Instruction Multiple Data)



## Shared-memory SIMD machines

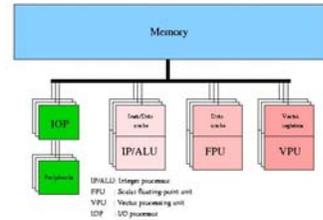


Diagrama de bloco de processador vetorial.

## Shared-memory SIMD machines

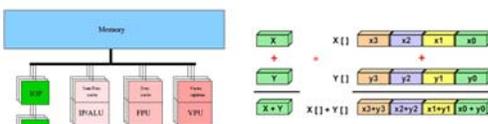


Diagrama de bloco de processador vetorial.

## Distributed-memory SIMD machines

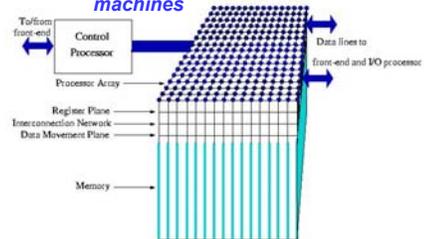


Diagrama de bloco genérico de uma máquina SIMD com memória compartilhada distribuída.

## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

Exemplo de computadores com a *arquitetura processor array* são as máquinas ILLIAC IV (Universidade de Illinois), Thinking Machine CM- 2 e MASP PAR MP-1216.



ILLIAC IV



Thinking Machines CM-2



MasPar

## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

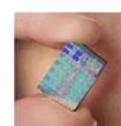
Exemplo de computadores com a *arquitetura vector pipeline* são as máquinas IBM 9000, Cray X-MP, Y-MP & C90, Fujitsu VP, NEC SX-2, Hitachi S820, ETA10. As GPUs, também, estão sob essa classificação



Cray X-MP



Cray Y-MP



Cell Processor (GPU)

## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

### MISD (Multiple Instruction Single Data)

Um conjunto de dados é colocado concorrente em múltiplas unidades de processamento. Cada UP opera de maneira independente via conjuntos independentes de instruções.

- Algumas utilizações de uma configuração MISD poderia ser:
- Filtros de múltiplas frequências operando um mesmo sinal;
  - Múltiplos algoritmos de criptografia tentando a quebra de uma mensagem codificada.

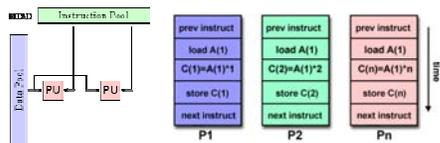
## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

### MISD (Multiple Instruction Single Data)

Não se tem conhecimento de arquitetura de máquinas comercial com múltiplas instruções trabalhando com um único conjunto de dados concorrente. Em 1971 um máquina denominada como C.mmp computer foi desenvolvida na universidade de Carnegie-Mellon.

## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

### MISD (Multiple Instruction Single Data)



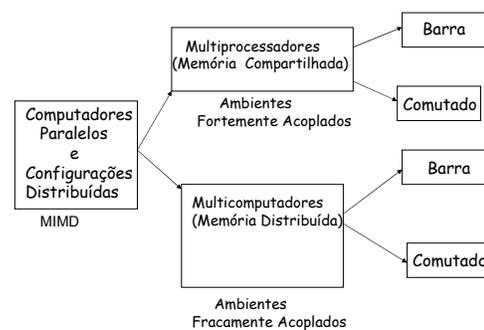
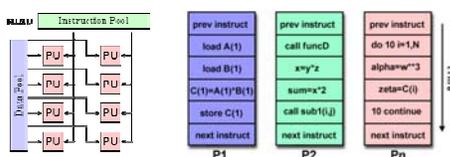
## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

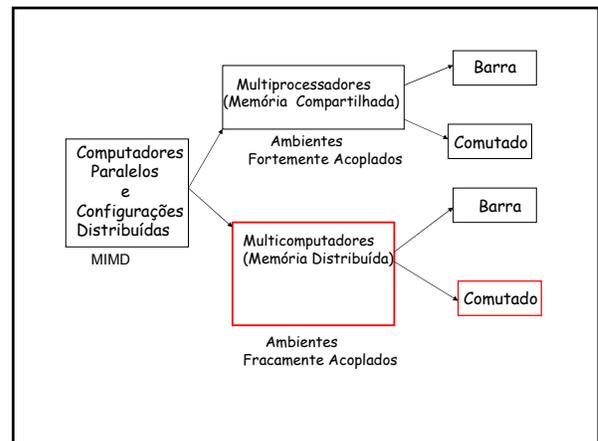
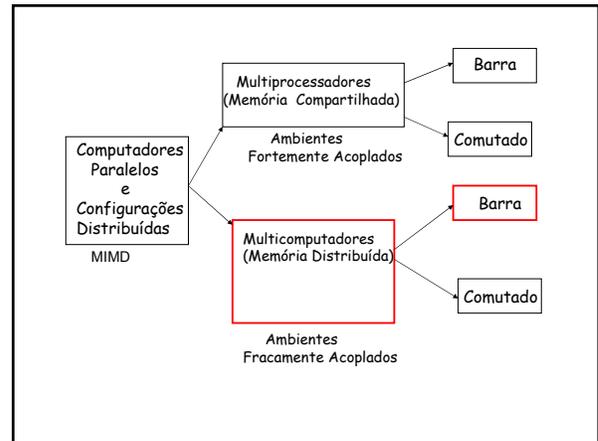
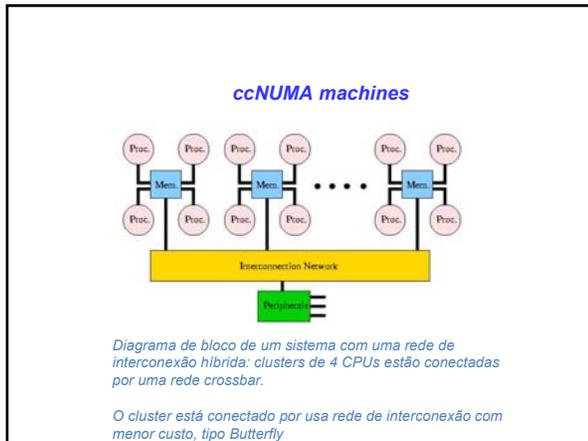
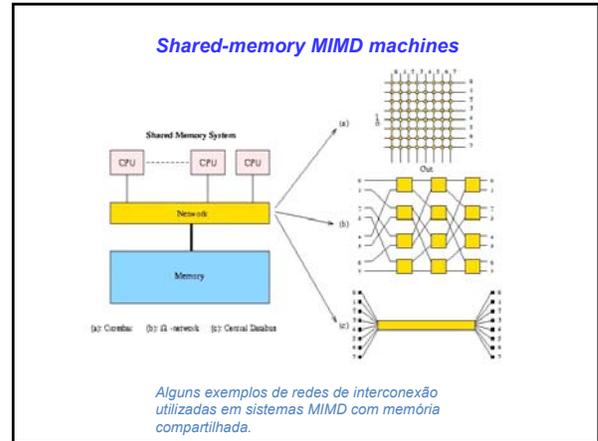
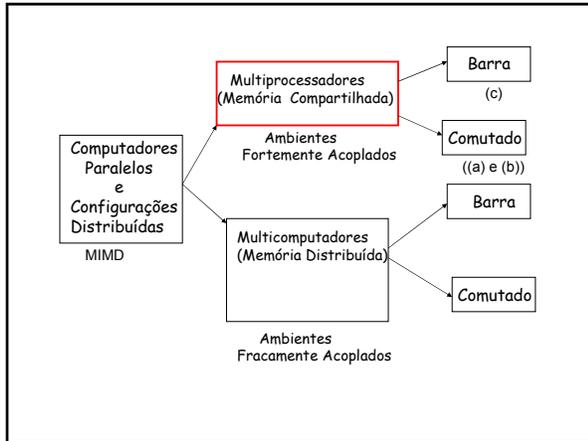
### MIMD (Multiple Instruction Multiple Data)

Arquiteturas sob esta classificação têm múltiplos processadores, cada qual podendo executar instruções independente dos demais.

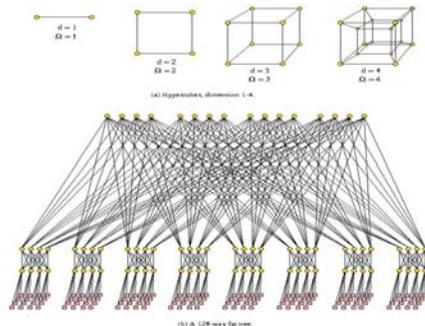
## Taxonomia de Arquitetura de Computadores Taxonomia de Flynn

### MIMD (Multiple Instruction Multiple Data)





### Distributed-memory MIMD machines



Exemplos de redes de interconexão utilizadas em máquinas com memória distribuída.

### Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- **Redes de Interconexão;**
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

### Redes de Interconexão

Rede de Interconexão (Interconnection Networks) são redes de altíssima taxa de transferência, projetadas para interconectar processadores e memórias numa arquitetura paralela. A ordem de grandeza da largura de banda é de Gbytes/sec e o retardo na casa dos  $n$  segundos.

### Redes de Interconexão

As métricas utilizadas para comparação das redes de interconexão são:

- Conectividade: Nós de uma rede e os enlaces podem apresentar falhas e devem ser removidos para reparo. A rede deve continuar funcionando com sua capacidade reduzida. O parâmetro conectividade indica a capacidade de flexibilidade da rede continuar a funcionar sob essas condições. A conectividade pode ser entendida como o número mínimo de redes e nós que ao falharem dividem a rede em redes disjuntas. Quanto maior a conectividade, melhor e a capacidade da rede em tratar com falhas;
- Diâmetro: representa a distância máxima inter-nó, ou seja, número máximo de enlaces que devem ser percorridos para envio de uma mensagem para qualquer nó ao longo do menor caminho. Quanto menor for o diâmetro, menor deverá ser o tempo de envio de uma mensagem de um nó para o nó mais distante;

### Redes de Interconexão

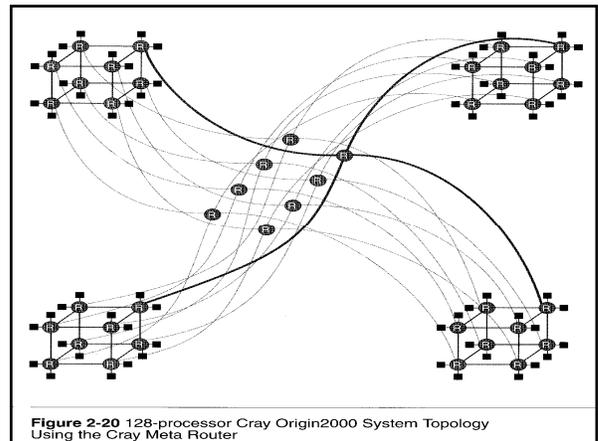
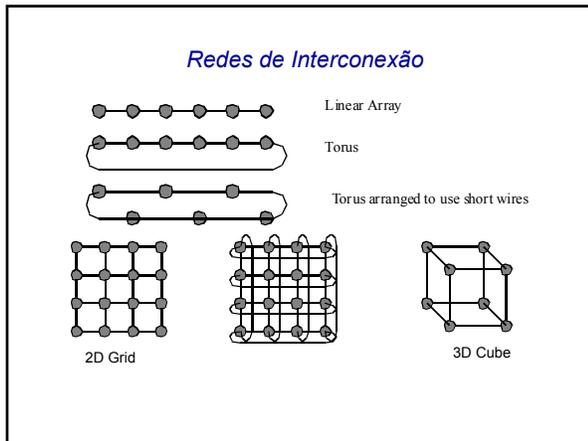
As métricas utilizadas para comparação das redes de interconexão são:

- Limitação: é a medida de congestionamento na rede. O cálculo é realizado da seguinte forma:
  - A rede é particionada em dois grupos de nós A e B, aonde o número em cada grupo é representado como  $N_a$  e  $N_b$  e  $N_b \ll N_a$ . Efetua-se a contagem  $I$  de números de interconexões entre A e B. Acha-se o valor máximo de  $N_b/I$  para todas as partições da rede. Esse número é a limitação da rede. A ideia é que se a limitação é alta ( $N_b > I$ ), então se  $N_b$  quiser enviar mensagem para  $N_a$ , o congestionamento vai ser alto, uma vez que teremos poucos enlaces para muitos processadores;
- Expansão: uma rede deve poder se expandir, criando maiores e mais potentes multicomputadores, através do acréscimo de mais nós. É desejável que o crescimento seja possível por pequenos incrementos.

### Redes de Interconexão

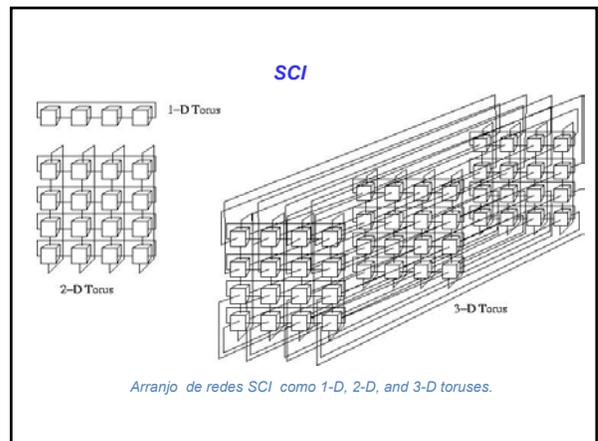
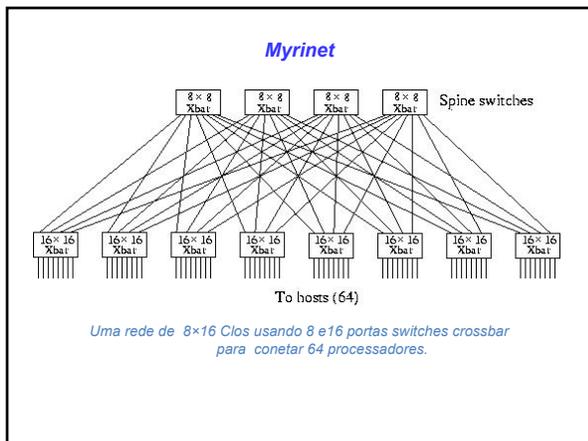
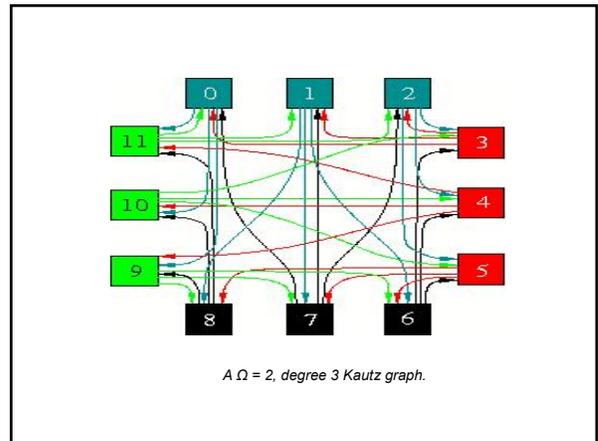
Redes de Interconexão podem ser configuradas de uma forma estática ou dinâmica. Em outras palavras, totalmente Interligada ou interligada dinamicamente.

- Topologias adotadas geralmente usadas são:
- linear array, ring, star, tree, nearest-neighbor mesh, systolic array, completely connected, 3-cube, 4-cube



### Redes de Interconexão

Network	Bandwidth GB/s	Latency $\mu$ s
Cray SeaStar2 (measured)	2.1	4.5
IBM (Infiniband) (measured)	1.2	4.5
SiCortex Kautz graph (stated)	2.0	1.0
SGI NumaLink (measured)	2.7	1.2
Infiniband (measured)	1.3	4.0
Infinipath (measured)	0.9	1.3
Myrinet 10-G (measured)	1.2	2.1
Quadrics QsNet <sup>®</sup> (measured)	0.9	2.7



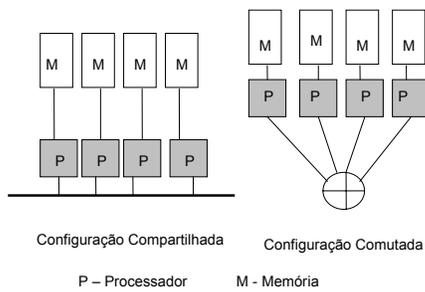
### Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- **Multicomputadores;**
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessadores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

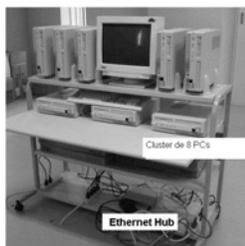
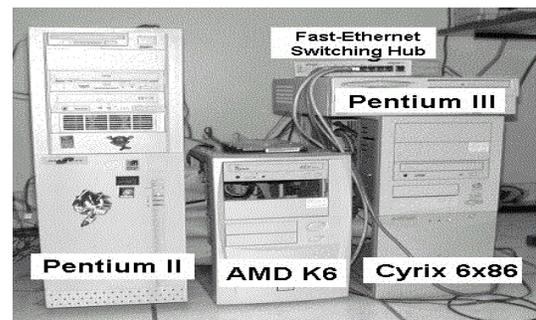
### Multicomputadores

Os computadores com arquitetura conhecida como *multicomputadores* são ambientes *fracamente acoplados*.

Em outras palavras, estas configurações caracterizadas por centenas (ou até milhares) de processadores têm suas próprias memórias locais.



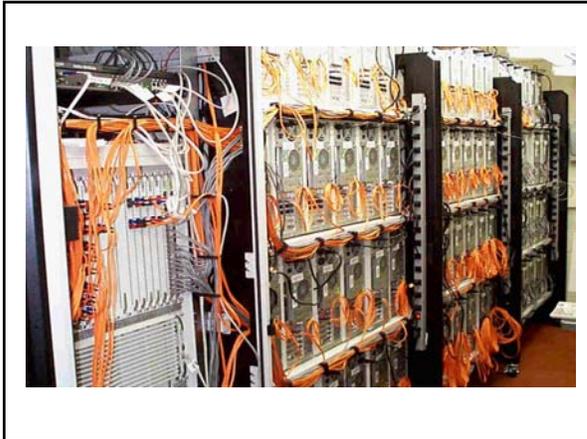
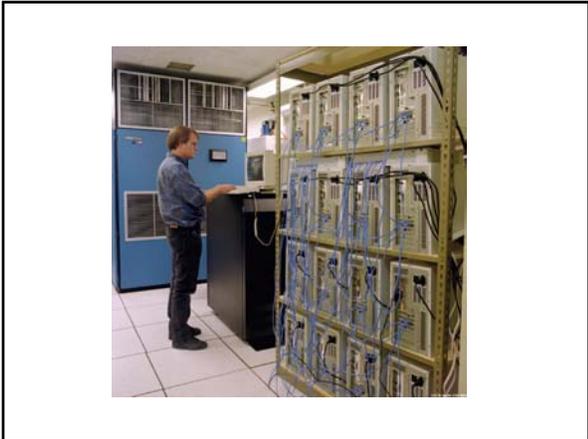
### Ambiente de "sucatas"



(a) Cluster dedicado



(b) Cluster não-dedicado

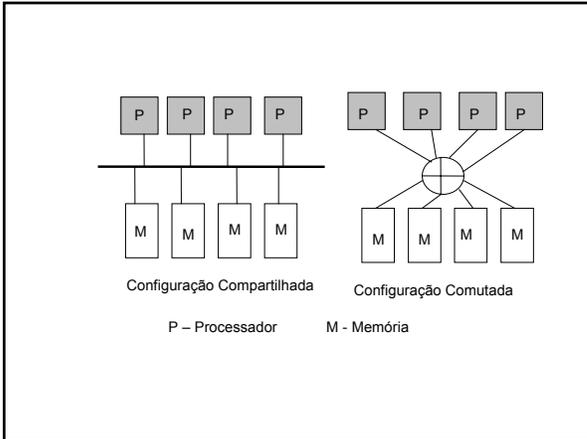


**Ementa**

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- **Multiprocessadores;**
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

**Multiprocessadores**

Estas arquiteturas são caracterizadas por vários processadores compartilhando uma única memória, ou um conjunto de memórias (não devemos nos esquecer que a memória de um computador pode fisicamente ser composta por *uma placa única* de memória ou composta por *conjunto* de placas de memórias).



### Multiprocessadores

A arquitetura dos multiprocessadores é conhecida como *fortemente acoplada*, uma vez que processadores e memória estão fortemente interligados através de seu sistema local de interconexão.

### Multiprocessadores

A arquitetura de um multiprocessador é caracterizada pelo compartilhamento global da memória pelos diversos processadores do ambiente. A escalabilidade em uma configuração multiprocessada varia entre alguns até centenas de processadores



(a) TX7 da NEC



(b) Altix da SGI

### Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- **Máquinas com Acesso Uniforme à Memória (UMA);**
- Multiprocessadores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

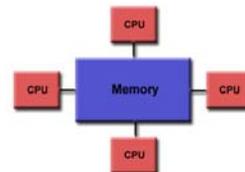
### Multiprocessadores (UMA)

A abordagem conhecida como UMA (*Uniform Memory Access*) é caracterizada por ter todos os elementos processadores com um acesso a memória.

Isto significa que todos tem a mesma taxa de transmissão e retardo.

### Multiprocessadores (UMA)

Na figura abaixo representa-se um exemplo teórico clássico de uma máquina com arquitetura NUMA:



### Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- **Multiprocessadores simétricos (SMP)**;
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

### Multiprocessadores Simétricos (SMP)

Os ambientes denominados como multiprocessadores simétricos (*Symmetric MultiProcessor - SMP*), são conhecidos como

*arquiteturas de compartilhamento total.*

Estas configurações são caracterizadas por até dezenas de processadores compartilhando todos os recursos computacionais disponíveis e executando um único sistema operacional.

### Multiprocessadores Simétricos (SMP)

Os processadores são considerados simétricos, uma vez que têm os mesmos custos para acesso à memória.

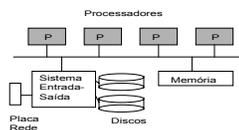
Todos, por exemplo, possuem acesso igual à memória e a qualquer dispositivo conectado no sistema de entrada e saída.

### Multiprocessadores Simétricos (SMP)

Um exemplo clássico de uma configuração SMP é ilustrado através da próxima figura.

Esta figura pode auxiliar o leitor a compreender melhor o conceito de uma máquina SMP.

### Multiprocessadores Simétricos (SMP)



*Configuração clássica de uma arquitetura SMP.*

### Multiprocessadores Simétricos (SMP)

Observe que a configuração é caracterizada por vários processadores compartilhando uma única memória e um único sistema de entrada e saída.

Um fator particular da configuração é não possuir múltiplas memórias e nem tão pouco múltiplos sistemas de entrada e saída, mas apenas múltiplos processadores.

### *Multiprocessadores Simétricos (SMP)*

A utilização de configurações SMP é mais popular do que se possa imaginar.

Exemplos comerciais que empregam esta abordagem são os servidores de fabricantes como a Compaq, IBM, Dell e HP.

Estas máquinas são usualmente denominadas pelos fabricantes como servidores de pequeno porte, pois possuem até dezena de processadores.

### *Multiprocessadores Simétricos (SMP)*

Os pequenos servidores visam melhorar o desempenho de aplicações por intermédio do compartilhamento do código com uma memória única utilizando mais intensamente os processadores disponíveis.

Uma aplicação exemplo seria um sistema de banco de dados. Com uma maior capacidade de processamento, uma máquina SMP é ideal para prover uma maior rapidez nas consultas e atualizações nos bancos de dados.

### *Multiprocessadores Simétricos (SMP)*

Na literatura sobre arquitetura de computadores alguns autores têm uma definição mais relaxada sobre os ambientes SMP.

Existem algumas configurações onde encontramos processadores que possuem suas próprias memórias, ou ainda um determinado processador com acesso exclusivo ao sistema de entrada e saída.

### *Multiprocessadores Simétricos (SMP)*

Sob outros pontos de vista, estas variações de configuração descaracterizam uma abordagem SMP pura, assim não consideramos como tal.

Estes ambientes podem não garantir o acesso uniforme com os mesmos custos à memória.

### *Multiprocessadores Simétricos (SMP)*

A arquitetura SMP sofre uma degradação à medida que o número de elementos que desejam se comunicar cresce na configuração.

Um outro paradigma possível de ser empregado é o uso de uma interconexão comutada.

A utilização de comutadores (switches) como elemento de interligação entre processadores e memória não escalam bem.

Existe um aumento no custo de ordem quadrática em relação ao número de portas que são acrescentadas à configuração.

### *Multiprocessadores Simétricos (SMP)*

Existem soluções de interconexão entre processadores e memória que podem agregar largura de banda à configuração, no entanto o custo não se torna proibitivo.

Naturalmente, nestas soluções devemos considerar um certo retardo que deverá ser acrescido para o acesso à memória.

## Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessadores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

## Multiprocessadores (NUMA)

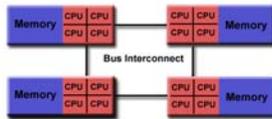
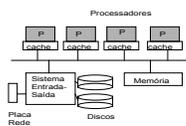
Uma arquitetura denominada de:

### Acesso Não-Uniforme à Memória (Non-Uniform Memory Access - NUMA)

é conhecida por sua característica de poder escalar até centenas de processadores.

Nss próximas figuras encontram-se configurações convencionais do tipo NUMA.

## Multiprocessadores (NUMA)



## Multiprocessadores (NUMA)

Um fator interessante das máquinas NUMA é que elas preservam o modelo de programação simples de uma configuração SMP.

Neste modelo de programação, processadores podem compartilhar os mesmos dados que estão armazenados em uma memória global.

## Multiprocessadores (NUMA)

Por outro lado, em uma configuração do tipo NUMA é reconhecido que existe um retardo de acesso uniforme a todas as regiões de memória.

Em outras palavras, é de conhecimento do programador que existe um retardo para acesso a determinadas partes da memória por determinados processadores dependendo de suas localizações.

## Multiprocessadores (NUMA)

O controlador da memória local pode decidir se uma determinada operação deve ser realizada na memória local, ou se uma transação de mensagem deve ser realizada com um controlador remoto.

O acesso à memória local é mais rápido, quando comparado ao acesso a uma memória remota.

Desta forma, é comum que o acesso à memória local seja efetuado às partes de código dos processos e que dados compartilhados sejam acessados em memória global.

### Multiprocessadores (NUMA)

O controlador da memória local pode decidir se uma determinada operação deve ser realizada na memória local, ou se uma transação de mensagem deve ser realizada com um controlador remoto.

O acesso à memória local é mais rápido, quando comparado ao acesso a uma memória remota.

Desta forma, é comum que o acesso à memória local seja efetuado às partes de código dos processos e que dados compartilhados sejam acessados em memória global.

### Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessadores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

### Multiprocessadores (ccNUMA)

A implementação prática de uma máquina com arquitetura NUMA é conhecida como máquina com:

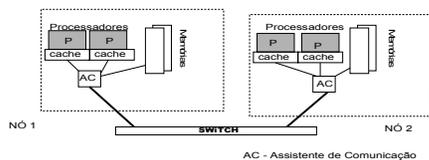
Acesso Não-Uniforme à Memória com Coerência de Cache (Cache Coherence Non-Uniform Memory Access – ccNUMA).

### Multiprocessadores (ccNUMA)

Máquinas com esta abordagem são configurações escaláveis de multiprocessadores, com apresenta a próxima figura.

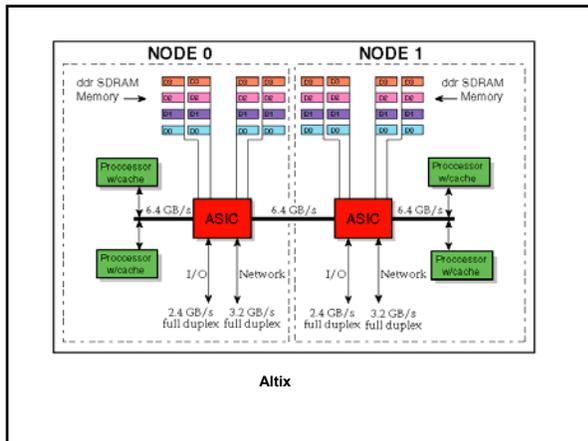
Aplicações tais como os serviços de Web, banco de dados, processamento de sinal, CRM e ERP são aplicações candidatas a serem utilizadas em configurações ccNUMA.

### Multiprocessadores (ccNUMA)



### Multiprocessadores (ccNUMA)

A seguir apresentamos alguns exemplos comerciais de máquinas com característica ccNUMA.



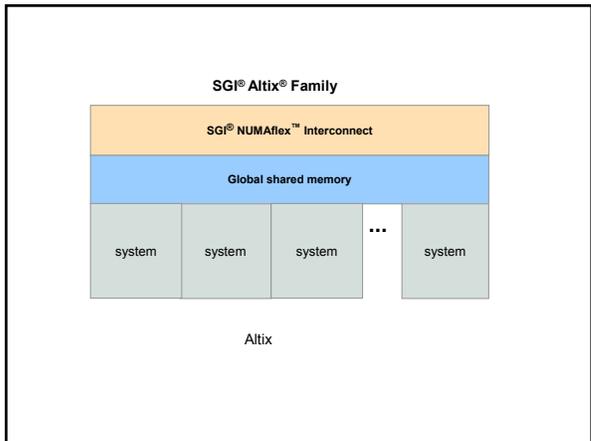
- Multiprocessadores (ccNUMA)*
- O computador ilustrado na figura anterior tem as seguintes peculiaridades:
- Dois nós com quatro processadores Intel Itanium;
  - Em cada barramento existem dois processadores interligados com taxa de transmissão de 6.4 GB/segundo; quatro módulos de memória de 64 GB;
  - Os controladores de memória têm taxa de transmissão entre 8.51 e 10.2 GB/segundo;
  - O sistema de interconexão com taxa de transmissão de 6.4 GB/segundo;
  - Largura agregada de transmissão de entrada/saída de 4.8 GB/segundo.

*Multiprocessadores (ccNUMA)*

A visão de um usuário da configuração é ilustrada na próxima figura.

É interessante observar que para uma visão de alto nível existe a abstração de uma máquina com vários Processadores compartilhando uma única memória.

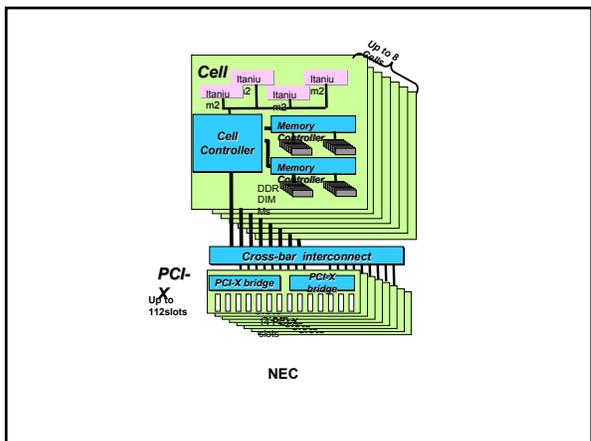
Todavia, sabemos que não existe realmente uma única memória, mas várias trabalhando como se fosse uma única.



*Multiprocessadores (ccNUMA)*

O exemplo apresentado a seguir é da NEC-HP, máquina conhecida como TX-7.

Esse multiprocessador é um exemplo clássico de ccNuma. Verifique a teoria que comentamos sobre uma configuração ccNuma e observe a máquina da próxima figura.



### Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- **Processadores Massivamente Paralelos (MPP)**;
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.

### Processadores Massivamente Paralelos (MPP)

As máquinas com configuração massivamente paralelas (*Massively Parallel Processors – MPP*), são conhecidas como arquiteturas fracamente acopladas. Computadores sob este paradigma são classificados usualmente como multicomputadores.

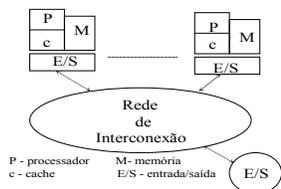
### Processadores Massivamente Paralelos (MPP)

Deve-se entender que um MPP pode ser composto também por um conjunto de multiprocessadores, onde cada multiprocessador é um nó de uma configuração do tipo multicomputador.

### Processadores Massivamente Paralelos (MPP)

Apresentamos um diagrama genérico, na próxima figura, de uma configuração de computador MPP. Interessante observar que a ordem de grandeza dos nós é de cerca de milhares.

### Processadores Massivamente Paralelos (MPP)



### Processadores Massivamente Paralelos (MPP)

Computadores com a arquitetura MPP são caracterizados por milhares de nós interligados por dispositivos de interconexão de alta velocidade.

Cada nó pode ser composto por um ou mais processadores, possuindo cache e memória locais.

### *Processadores Massivamente Paralelos (MPP)*

Uma outra característica da arquitetura é que cada *nó* possui sua própria cópia de sistema operacional, onde as aplicações executam localmente e se comunicam através de pacotes de troca de mensagem, tais como

- MPI (Message Passing Interface)
- PVM (Parallel Virtual Machine) .

### *Ementa*

- *Arquitetura de computadores modernos;*
- *Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);*
- *Redes de Interconexão;*
- *Multicomputadores;*
- *Multiprocessadores;*
- *Máquinas com Acesso Uniforme à Memória (UMA);*
- *Multiprocessores simétricos (SMP);*
- *Máquinas com Acesso Não Uniforme à Memória (NUMA);*
- *Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);*
- *Processadores Massivamente Paralelos (MPP);*
- **Sistemas Distribuídos;**
- *Clusters;*
- *Grids;*
- *Clouds.*

### *Sistemas Distribuídos*

Os sistemas distribuídos, sob o aspecto de arquitetura de máquinas para execução de aplicativos, devem ser vistos como configurações com grande poder de escala pela agregação dos computadores existentes nas redes convencionais.

### *Sistemas Distribuídos*

Nos ambientes distribuídos, a homogeneidade ou heterogeneidade de um conjunto de máquinas, onde cada qual possui sua arquitetura de software-hardware executando sua própria cópia de sistema operacional, permite a formação de interessantes configurações de SMPs, de MPPs, de *clusters* e *grids* computacionais.

### *Sistemas Distribuídos*

O termo *metacomputador (metacomputer)* é empregado muitas vezes como referência ao uso de sistemas distribuídos como um grande computador.

### *Sistemas Distribuídos*

A revista Scientific American publicou em Maio de 1997 uma afirmação dizendo que provavelmente o computador mais rápido existe atualmente é a Internet, ou um subconjunto de suas máquinas da Internet, agrupadas para a execução de uma aplicação.

## Sistemas Distribuídos

Aspectos tais como a segurança, o retardo de comunicação, a confiabilidade, a disponibilidade e a compatibilidade de versões de pacotes de software são alguns pontos a serem considerados com cautela em uma configuração distribuída.

## Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- **Clusters**;
- Grids;
- Clouds.

## Clusters

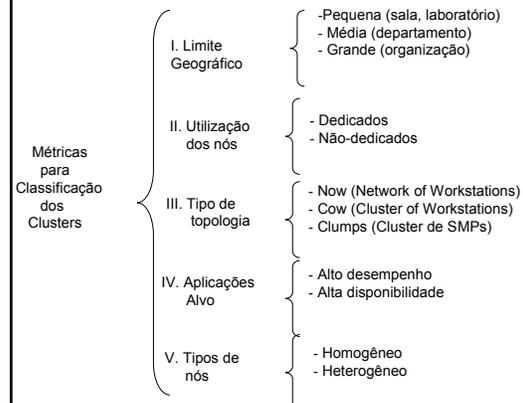
As configurações de *clusters*, em termos de arquiteturas computacionais, podem ser entendidas como uma agregação de computadores de uma forma dedicada (ou não) para a execução de aplicações específicas de uma organização.

## Clusters

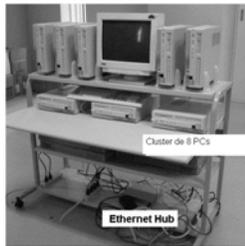
Ilustramos nos próximos exemplos um *cluster* dedicado, ou projetado para rodar exclusivamente as aplicações na configuração, e um outro que representa uma configuração não dedicada.

## Clusters

No segundo ambiente, além da execução de tarefas convencionais monoprocessadas pode ser utilizado como um *cluster* eventual para execução de aplicações que solicitem um maior desempenho computacional agregado.



### Clusters



Ambiente dedicado.

### Clusters



Ambiente não dedicado.

### Clusters

Os *clusters* (ou *agregados* com alguns autores se referem em português), de uma forma geral, são compostos por computadores do tipo IBM-PC com uma característica intrínseca de disponibilidade de uma grande quantidade de recursos (processadores, memórias e capacidade de armazenamento) pertencentes a uma única entidade (laboratório, departamento, filial ou empresa).

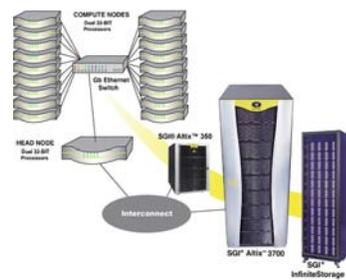
### Clusters

Deve -se entender que é possível projetarmos configurações de *clusters* levando-se em consideração um conjunto de máquinas multiprocessadas.

### Clusters

Na próxima figura apresentamos um *cluster* híbrido de máquinas multiprocessadas da SGI.

### Clusters



Cluster de máquinas multiprocessadas

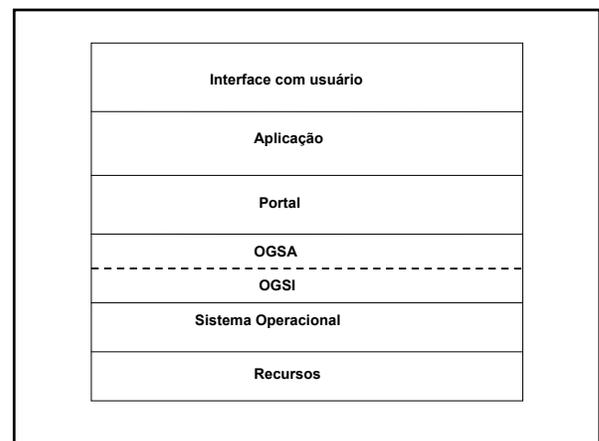
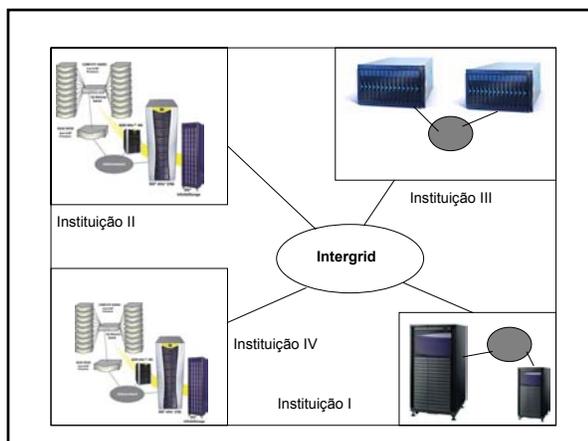
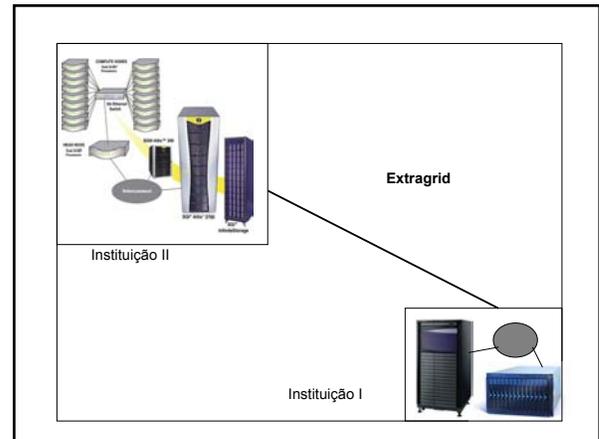
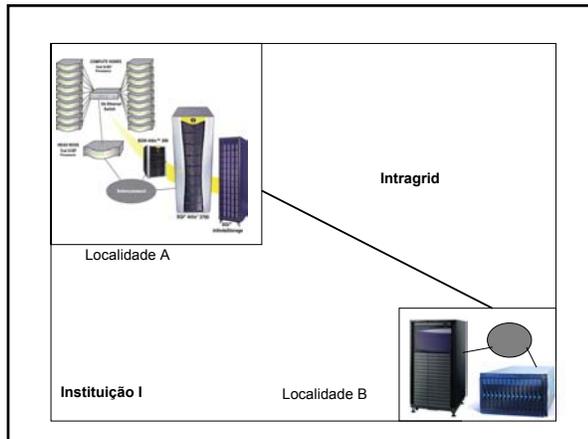
## Clusters

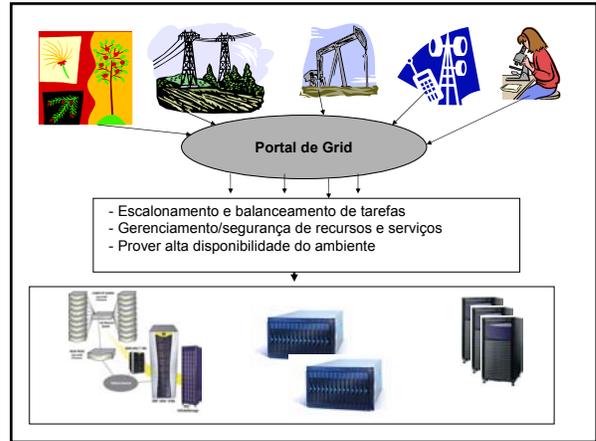
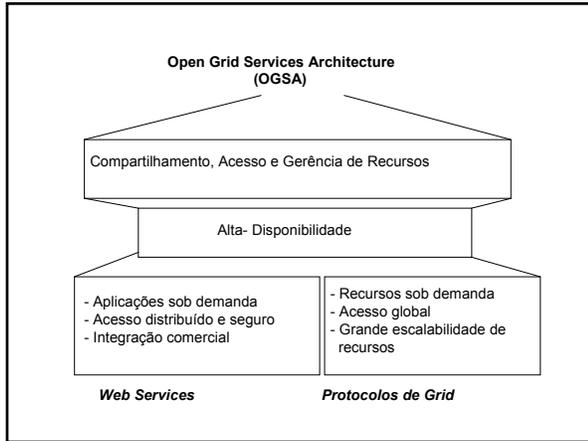
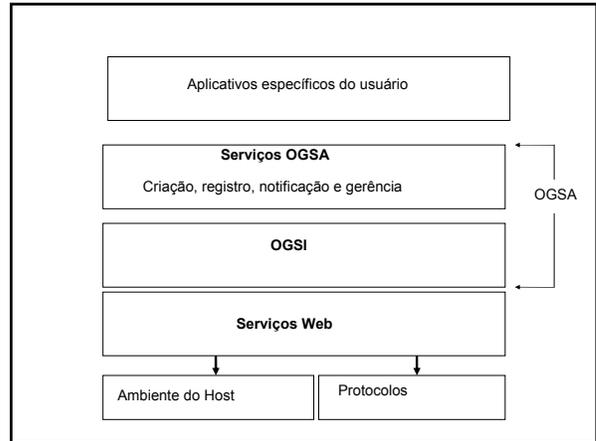
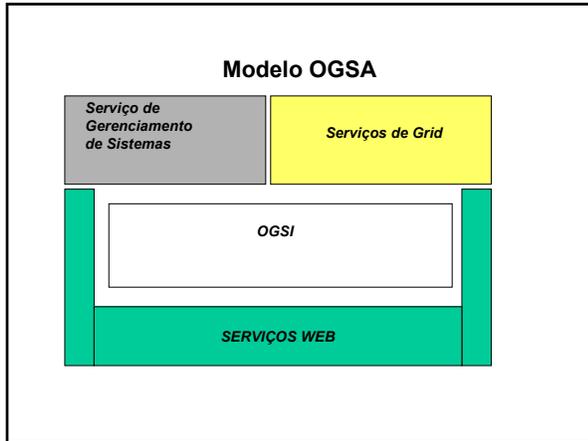
A escalabilidade é um fator diferencial dos ambientes de *cluster*, pois a configuração pode crescer à medida que mais recursos estiverem disponíveis.

A agregação de máquinas com configurações relativamente pequenas endereçam o paradigma de *pedir ajuda*.

## Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.





### Ementa

- Arquitetura de computadores modernos;
- Contextualização da taxonomia de Flynn (SISD, SIMD, MISD, MIMD);
- Redes de Interconexão;
- Multicomputadores;
- Multiprocessadores;
- Máquinas com Acesso Uniforme à Memória (UMA);
- Multiprocessadores simétricos (SMP);
- Máquinas com Acesso Não Uniforme à Memória (NUMA);
- Máquinas com Coerência de Cache e Acesso Não Uniforme à Memória (ccNUMA);
- Processadores Massivamente Paralelos (MPP);
- Sistemas Distribuídos;
- Clusters;
- Grids;
- Clouds.