



Subsistema de Entrada/Saída

Aleardo Manacero Jr.





Entrada/Saída

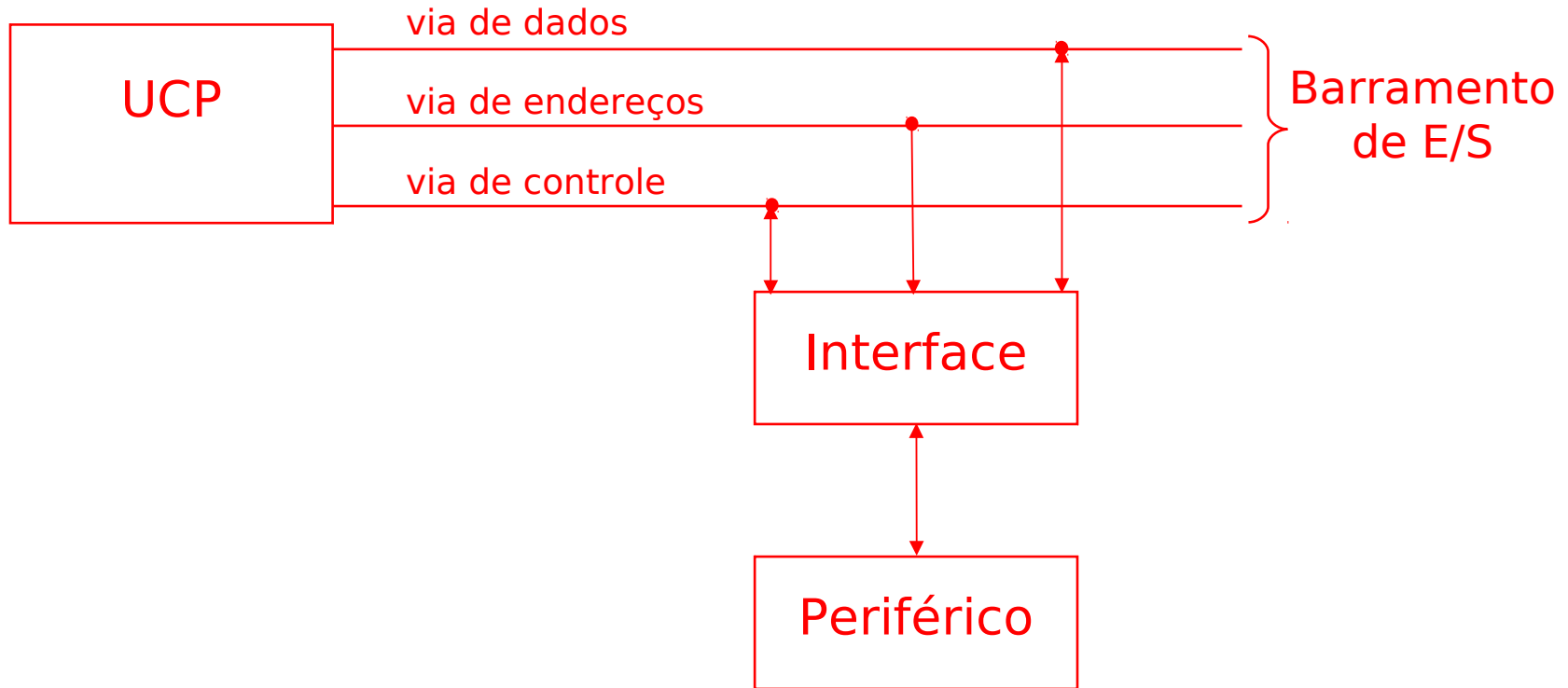


Conceitos

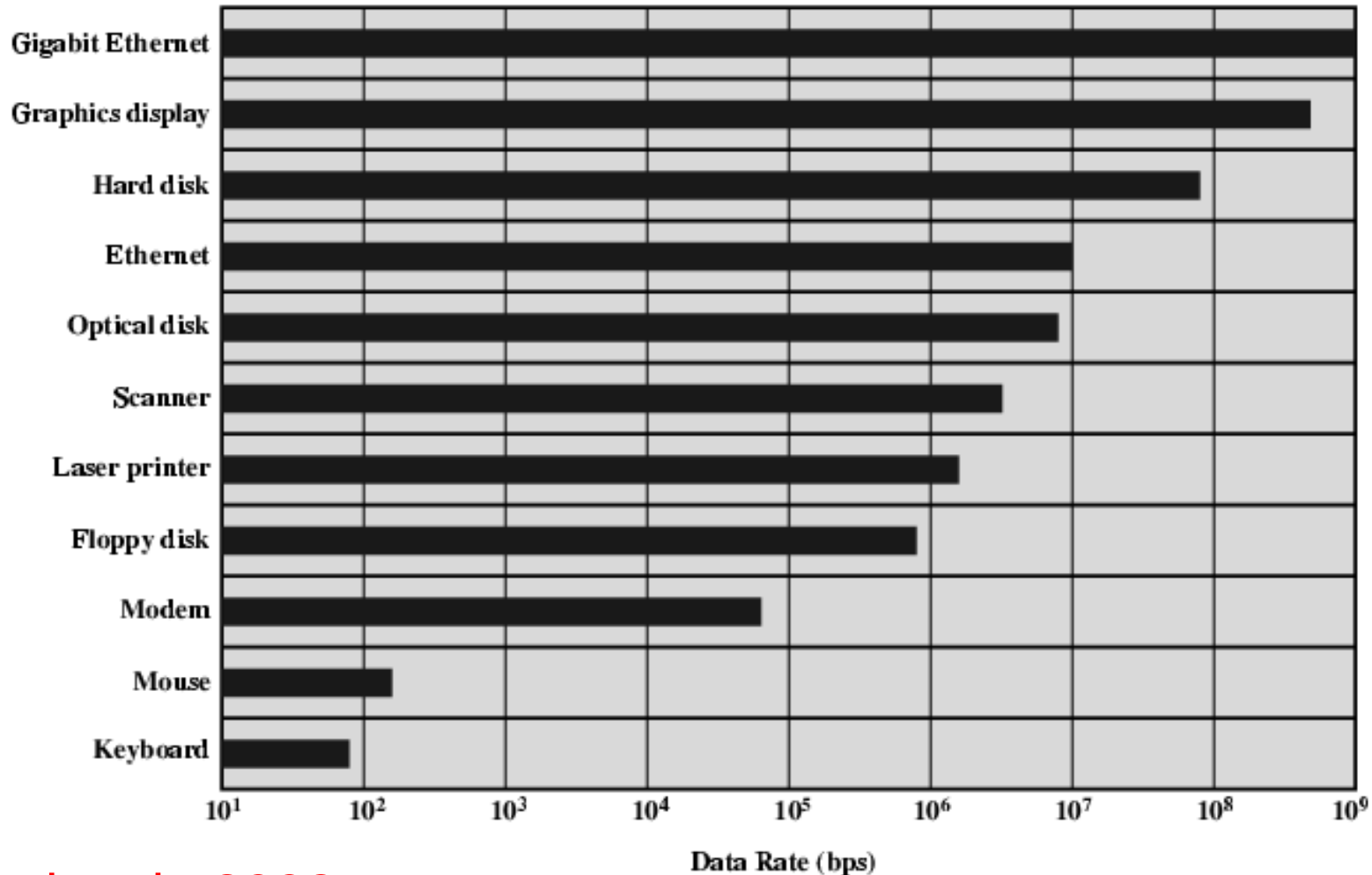


- Periférico
 - Entrada
 - Saída
 - Entrada/saída
- Interface
- Controlador
- Acionador
- Porta
- Barramento

Organização de E/S



Porque inserir uma interface?

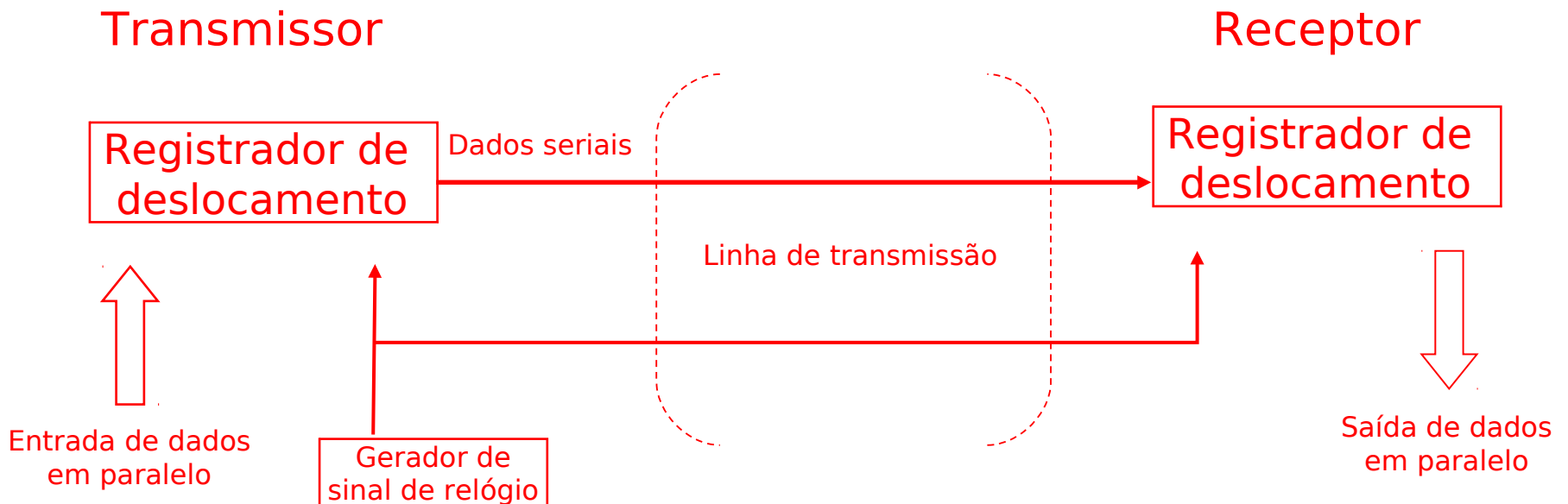


Dados de 2002

Transações de entrada/saída



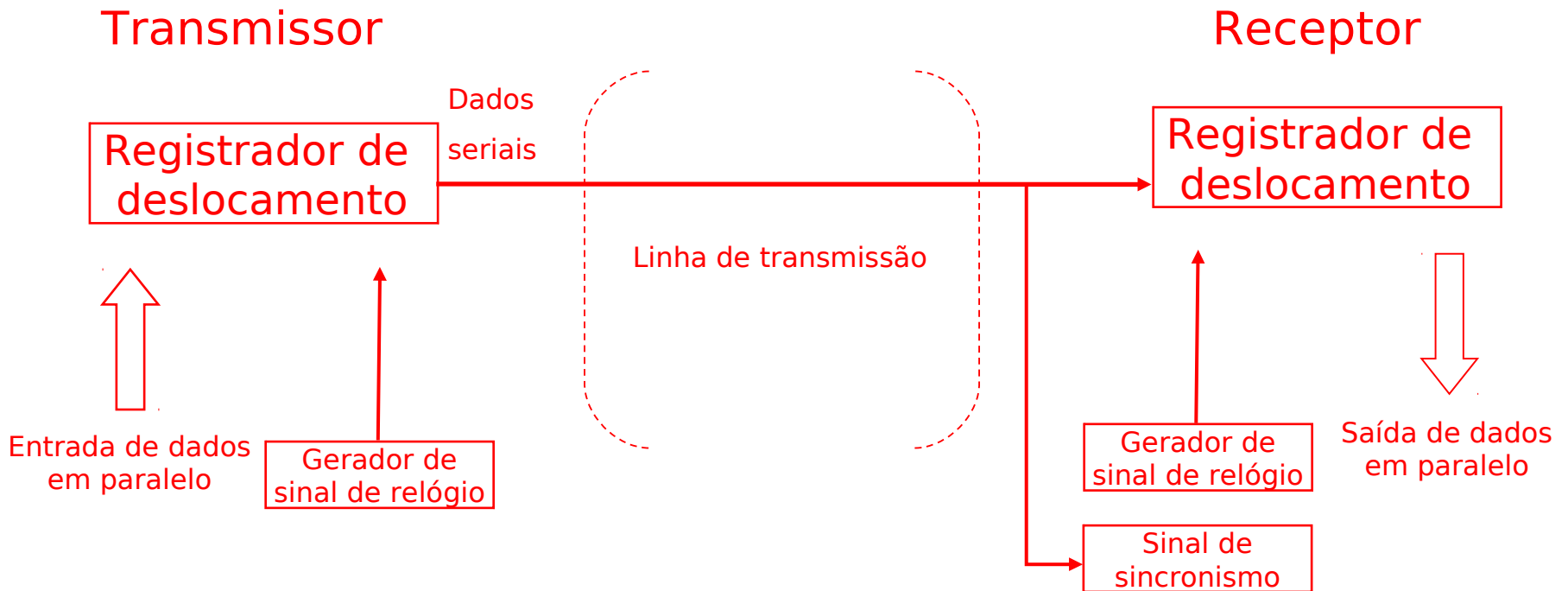
- Síncronas



Transações de entrada/saída



- Assíncronas



Transações de entrada/saída



- Malha aberta
 - Não há envio, pelo receptor, de sinais de reconhecimento ou do estado da informação recebida.
- Malha fechada
 - Utilização de protocolos de comunicação.

Transações de entrada/saída



Simplex



ou



Transações de entrada/saída



Semi duplex



Transações de entrada/saída



Duplex



e



Métodos de acesso a dispositivos



- Mapeamento em memória
 - Barramento único para memória e E/S
 - Espaço de endereçamento compartilhado
 - Instruções de memória = E/S
- Configuração isolada
 - Barramentos separados para memória e E/S
 - Espaços de endereçamento disjuntos
 - Instruções de memória $\neq$ E/S

Modos de transferência de dados

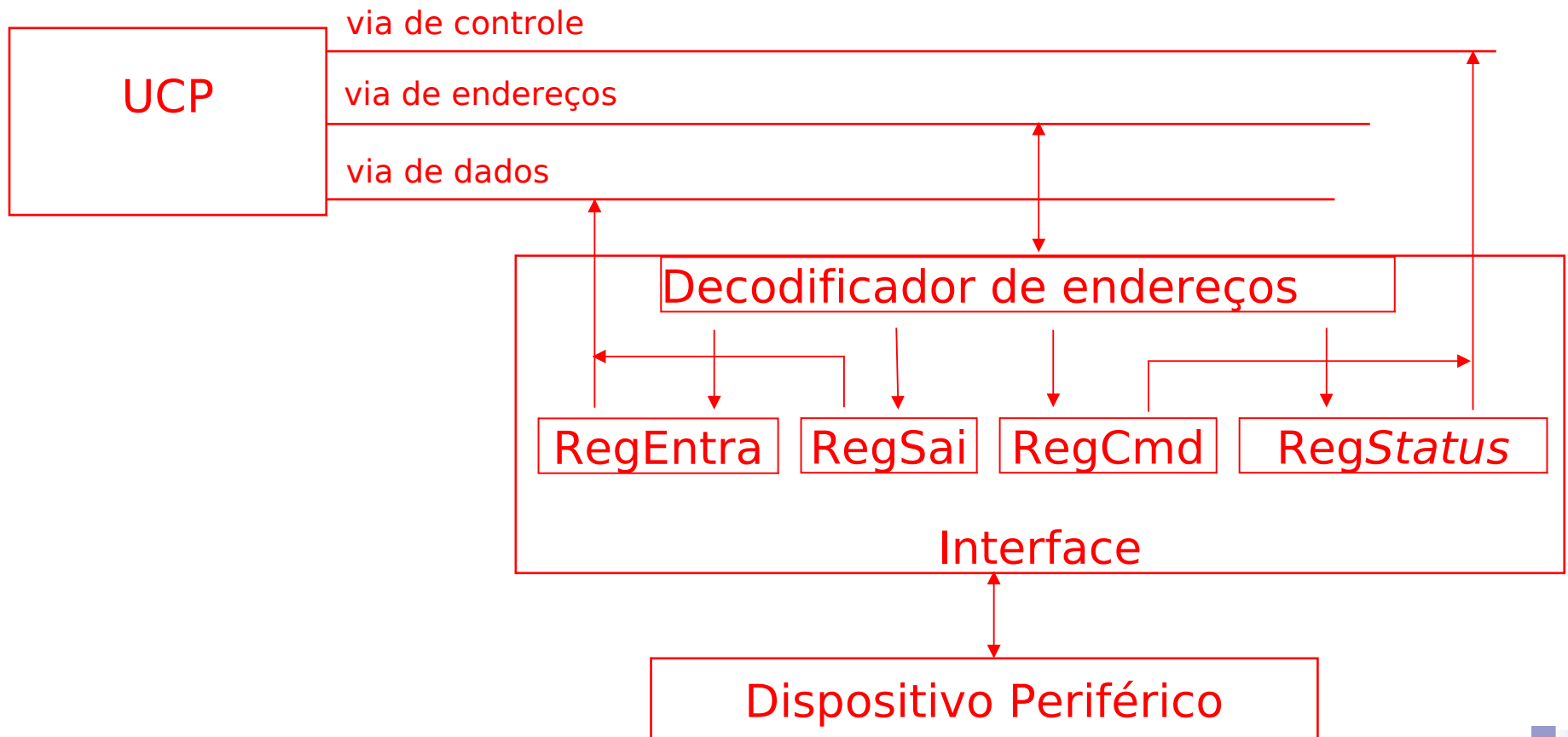


- Controle direto do processador
 - Teste de estado
 - Interrupções
- Acesso direto à memória
- Processadores de entrada e saída

Modos de transferência de dados



- Teste de estado $\equiv$ E/S Programada



Modos de transferência de dados



■ Interrupções

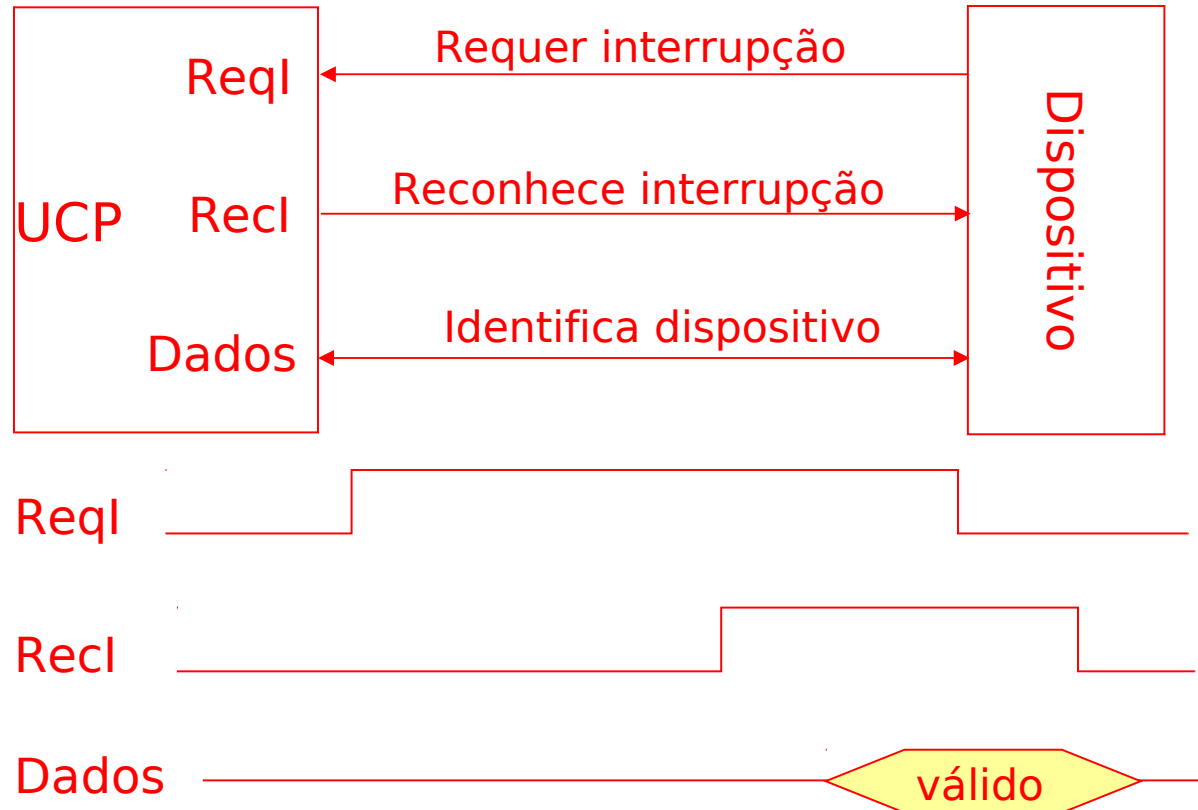
- ☐ Monitoramento do estado do dispositivo
- ☐ Solicitação de interrupção ao processador
- ☐ Salvamento do contexto
- ☐ Ativação da rotina correspondente
- ☐ Processamento de entrada ou saída
- ☐ Recuperação do contexto
- ☐ Retorno às atividades da UCP

Modos de transferência de dados



■ Interrupções

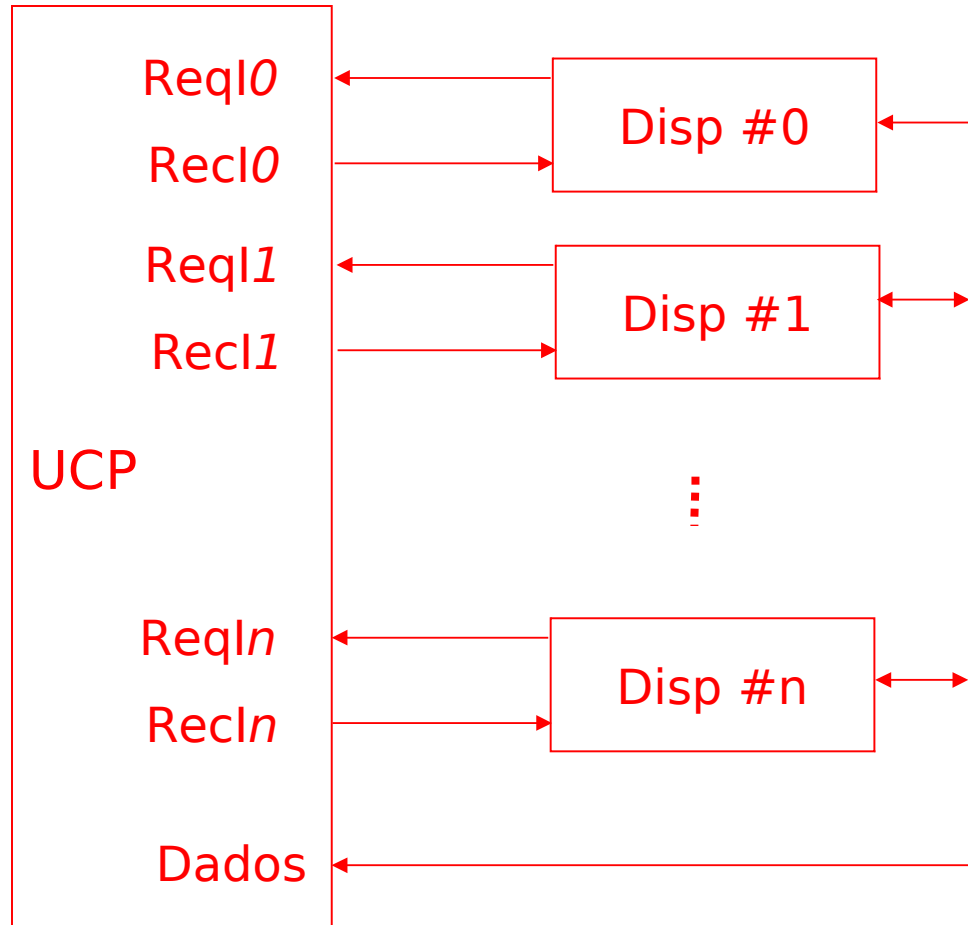
- Internas
- Por *software*
- Externas



Modos de transferência de dados



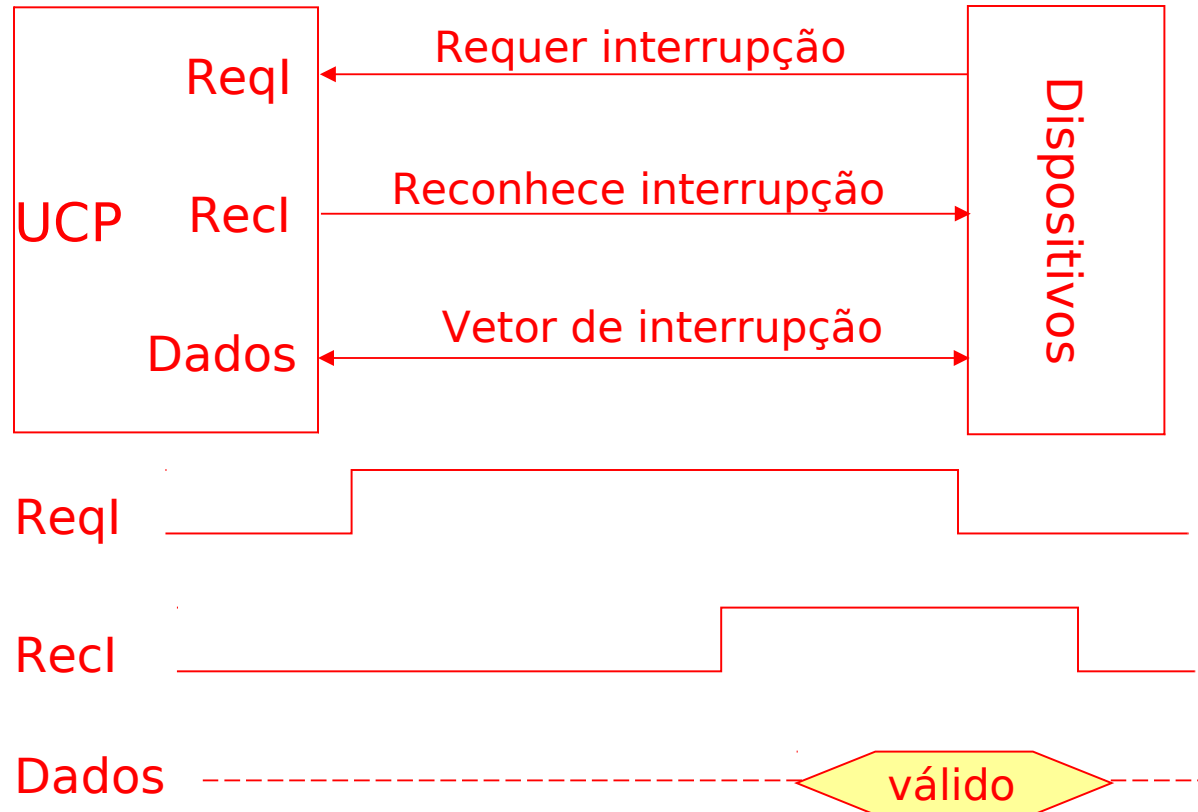
- Processamento de várias solicitações



Modos de transferência de dados



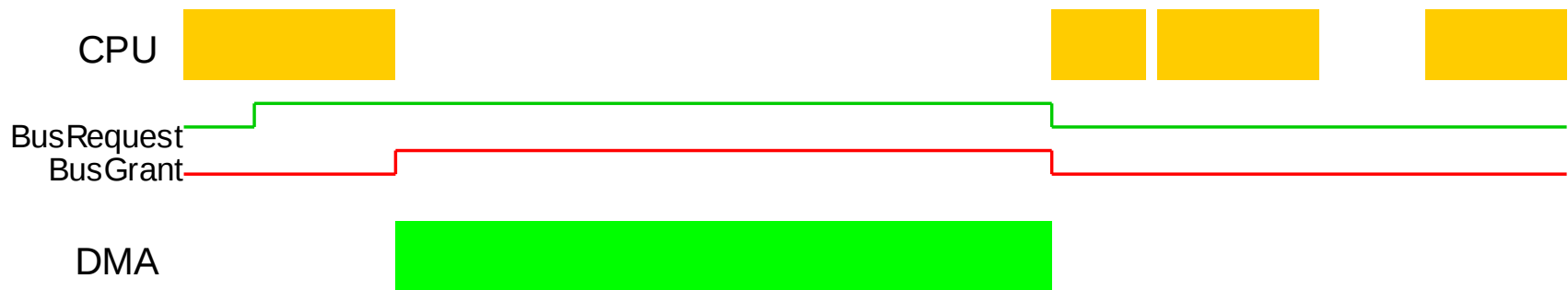
■ Interrupções vetorizadas



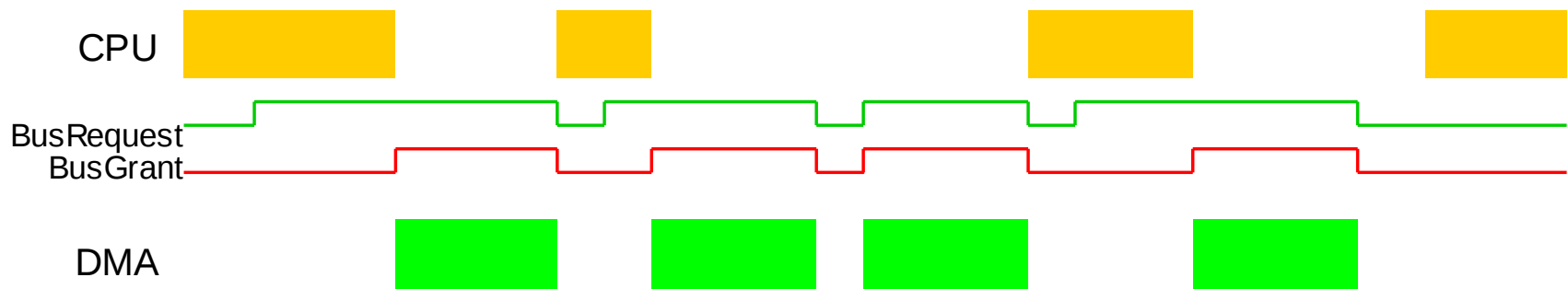
Modos de transferência de dados



■ Acesso Direto a Memória - DMA



(a) DMA transfer in one continuous burst

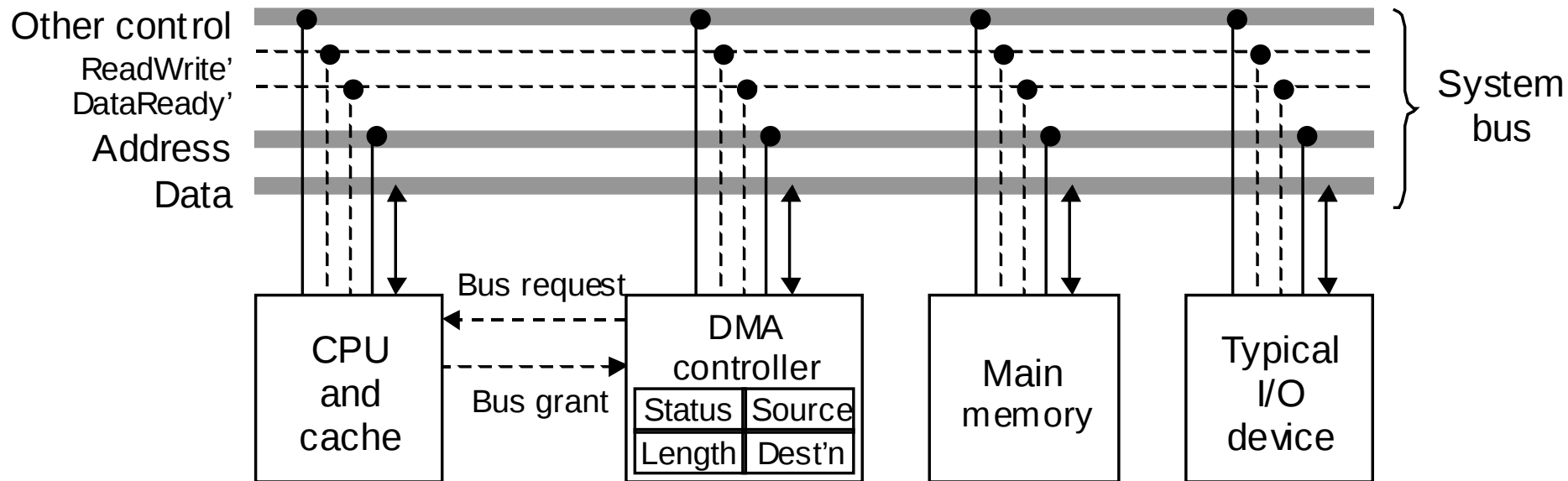


(b) DMA transfer in several shorter bursts

Modos de transferência de dados



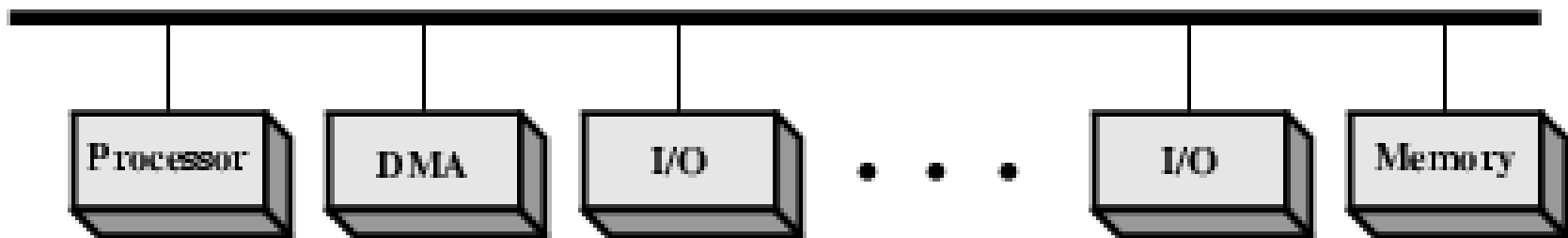
Acesso Direto a Memória - DMA



Modos de transferência de dados



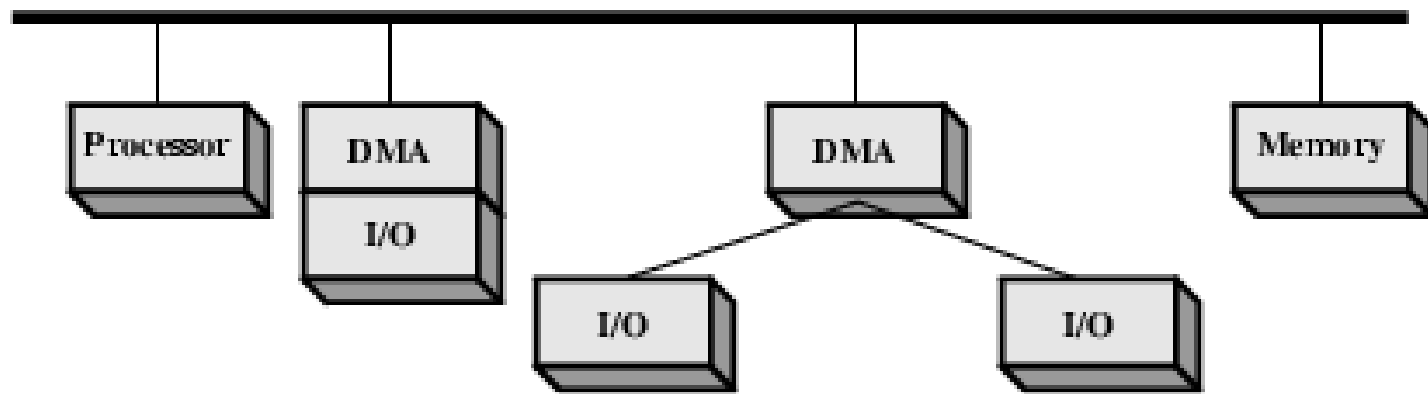
- Barramento único, DMA isolado
- Usa barramento duas vezes por transferência
 - I/O para DMA e depois entre DMA e memória
- CPU é suspensa duas vezes



Modos de transferência de dados



- Barramento único, DMA integrado
- Controlador pode suportar vários dispositivos
- Cada transferência usa o barramento uma vez



(b) Single-bus, Integrated DMA-I/O

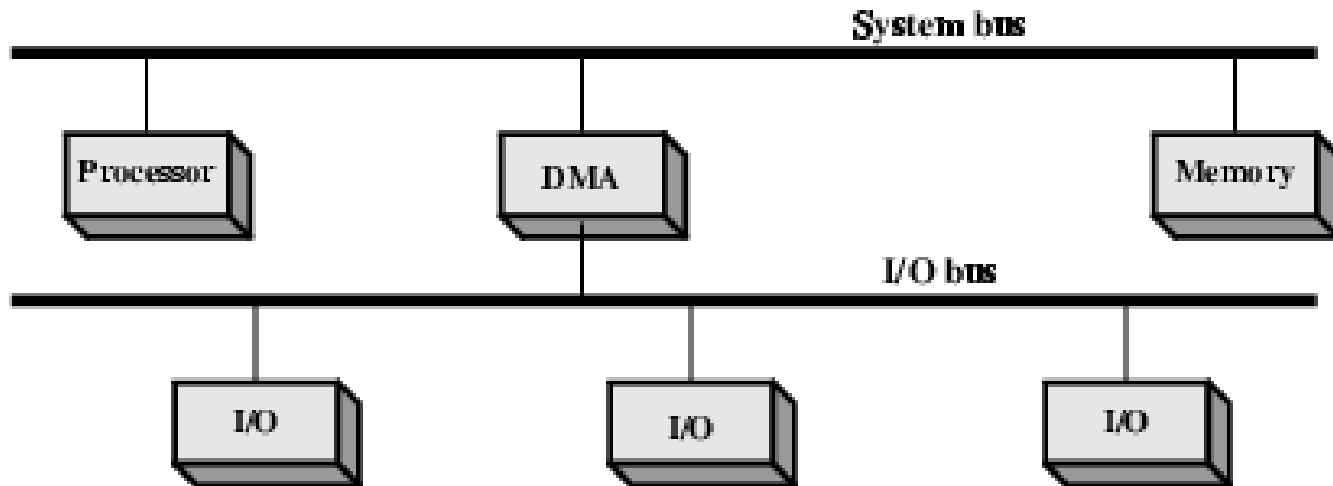
Modos de transferência de dados



- Barramento de E/S separado
- Barramento suporta todos dispositivos de DMA
- Cada transferência usa o barramento uma vez



■ CF



(c) I/O bus

Modos de transferência de dados



- Processadores de entrada e saída
 - Dispositivos de E/S estão mais sofisticados, como dispositivos gráficos p.ex
 - Permite a introdução de processadores dedicados para as operações de E/S
 - Esses controladores de E/S fazem toda a transferência, melhorando a velocidade e reduzindo a carga da CPU

Modos de transferência de dados

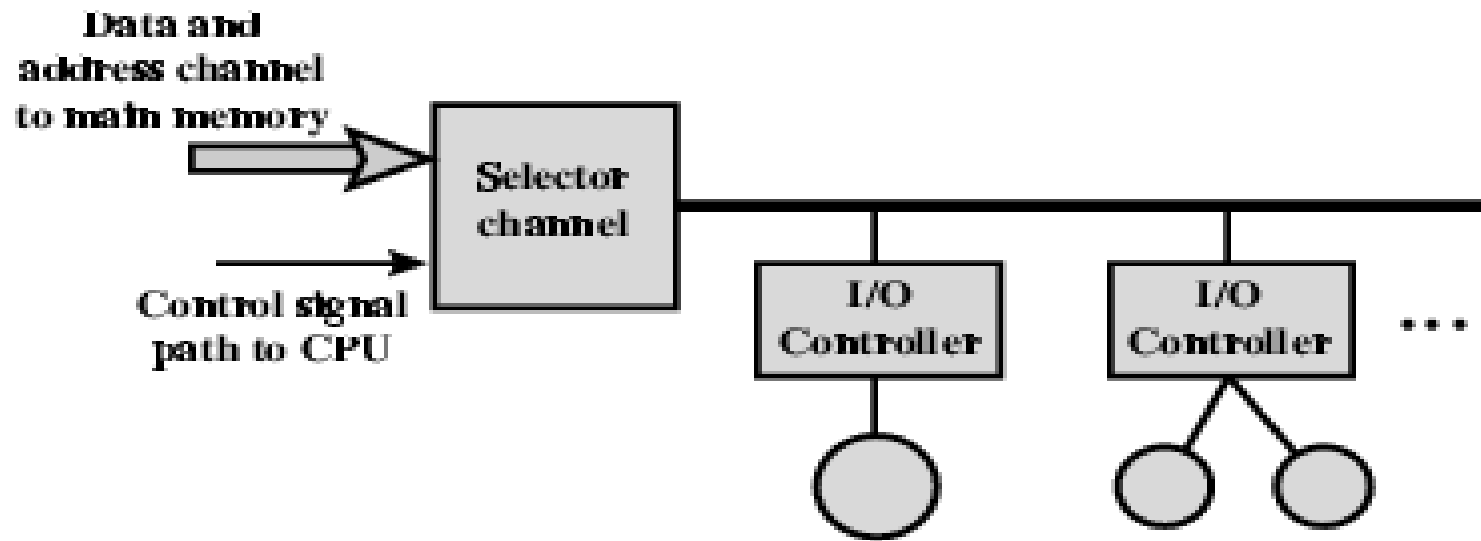


- Processadores de entrada e saída
 - Dispositivos de E/S estão mais sofisticados, como dispositivos gráficos p.ex
 - Permite a introdução de processadores dedicados para as operações de E/S
 - Esses controladores de E/S fazem toda a transferência, melhorando a velocidade e reduzindo a carga da CPU

Modos de transferência de dados



- Processadores de entrada e saída

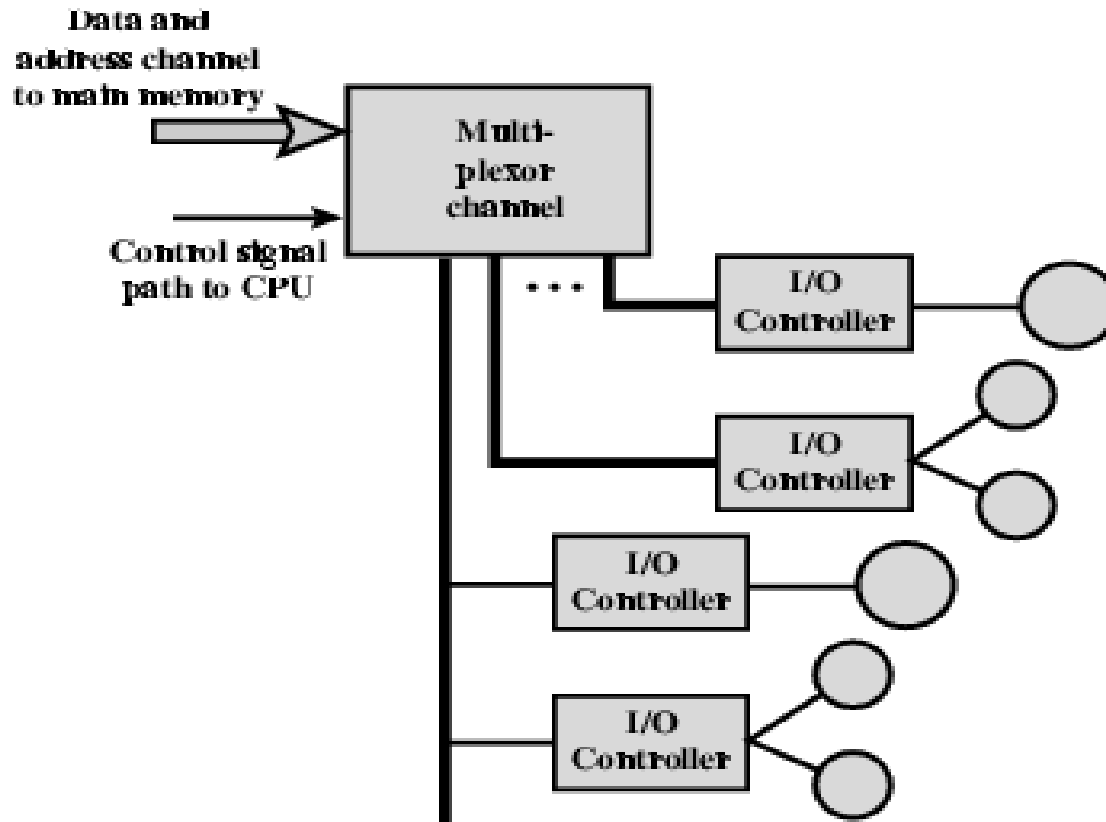


(a) Selector

Modos de transferência de dados



- Processadores de entrada e saída



(b) Multiplexor

Summary of four standard interface buses.

Attributes ↓	Name →	PCI	SCSI	FireWire	USB
Type of bus		Backplane	Parallel I/O	Serial I/O	Serial I/O
Standard designation		PCI	ANSI X3.131	IEEE 1394	USB 2.0
Typical application domain		System	Fast I/O	Fast I/O	Low-cost I/O
Bus width (data bits)		32-64	8-32	2	1
Peak bandwidth (MB/s)		133-512	5-40	12.5-50	0.2-15
Maximum number of devices		1024*	7-31 [#]	63	127 ^{\$}
Maximum span (m)		< 1	3-25	4.5-72 ^{\$}	5-30 ^{\$}
Arbitration method		Centralized	Self-select	Distributed	Daisy chain
Transceiver complexity or cost		High	Medium	Medium	Low





Elementos de Conexão



O que é conectividade?



- Define como os elementos de processamento (CPU e memória) de um sistema de alto desempenho serão interligados.

O que se deve considerar?



- Número de conexões que cada nó deve ter. Esse valor é o **GRAU do nó**.
- Caminho mínimo entre dois nós. O maior desses caminhos mínimos é definido como **DIÂMETRO da rede**.

O que se deve considerar?



- Quantos canais são necessários entre duas metades de uma rede. Isso é denominado **LARGURA DE BISSECÇÃO**.
- Formas de tráfego dos dados. Isso é definido através de uma **FUNÇÃO DE ROTEAMENTO**.

O que é conectividade?



- Define como os elementos de processamento (CPU e memória) de um sistema de alto desempenho serão interligados.

O que é conectividade?

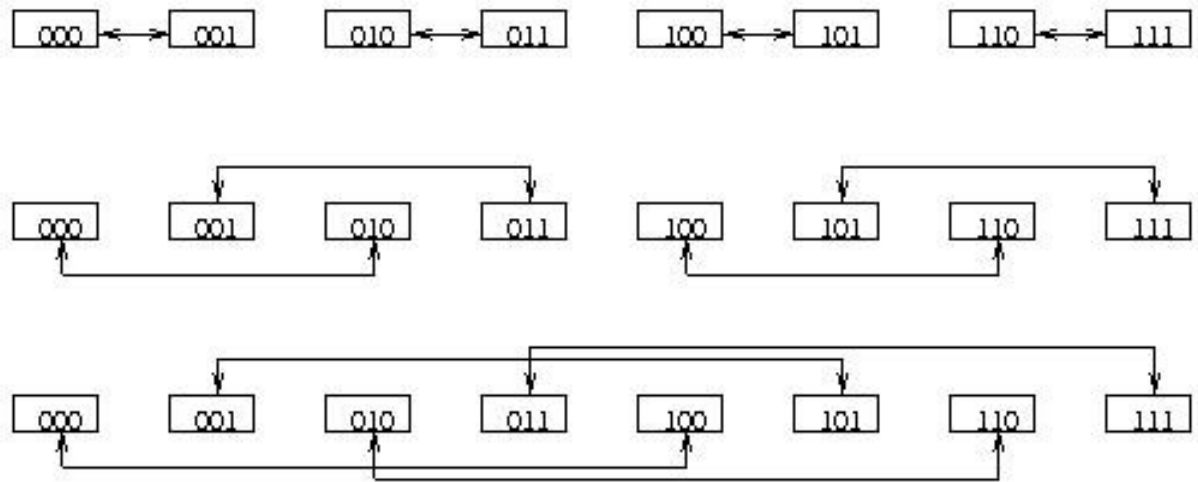
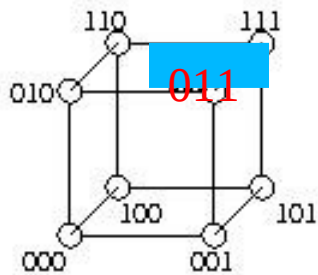


- Define como os elementos de processamento (CPU e memória) de um sistema de alto desempenho serão interligados.

Exemplos



I



I

Tipos de redes de conexão



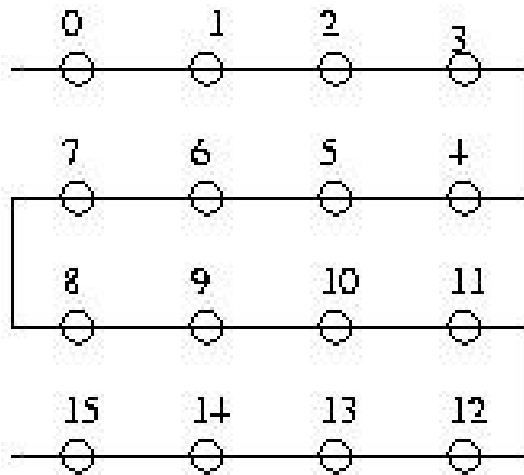
- Redes estáticas

Nessas redes as conexões entre os nós são fixas. As ligações são diretas entre cada dois pontos.

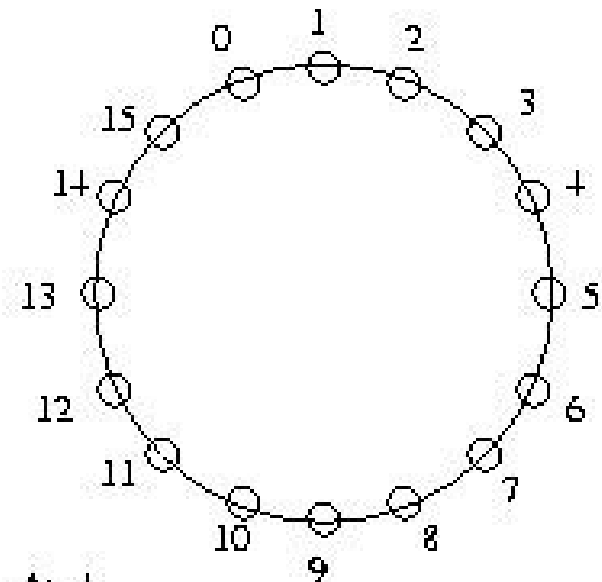
- Redes dinâmicas

Aqui as conexões variam com o tempo. Não existem ligações diretas entre dois nós.

Exemplos de redes estáticas

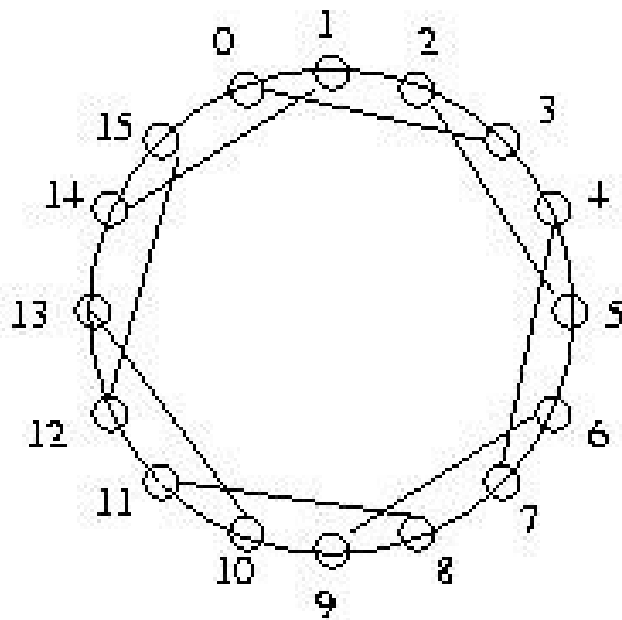


Array linear

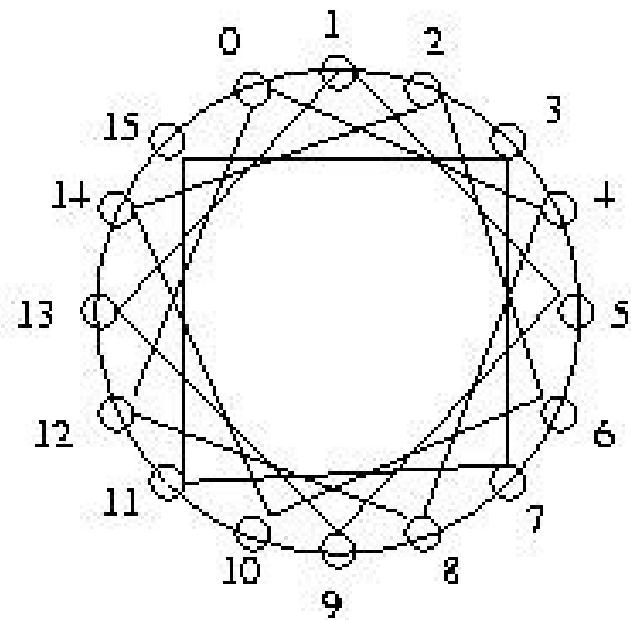


Anel

Exemplos de redes estáticas

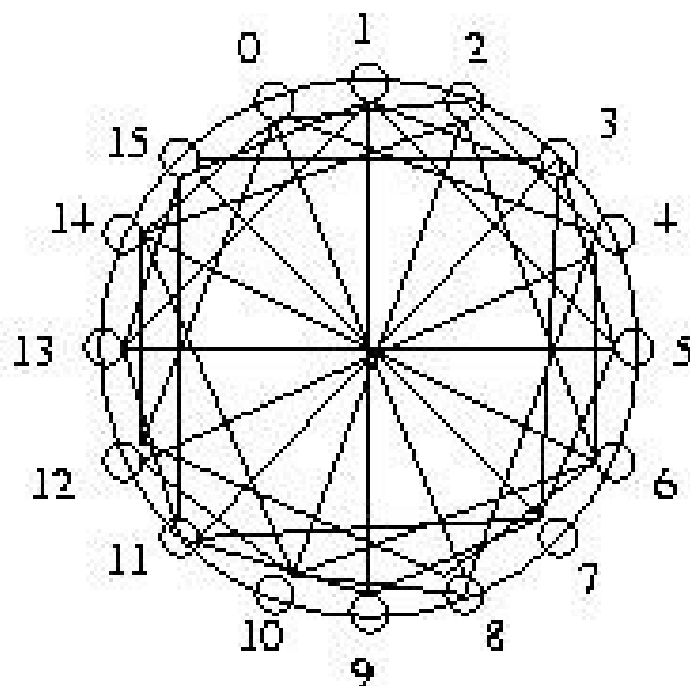


Anel corda 3

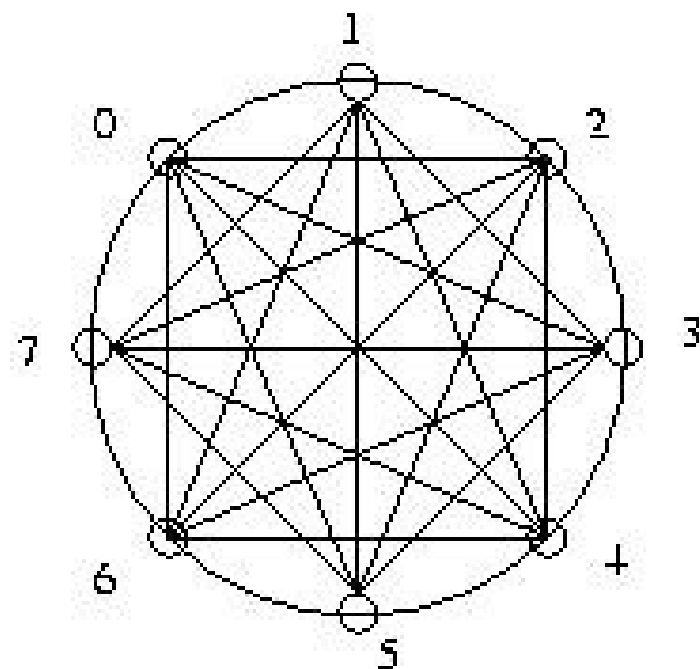


Anel corda 4

Exemplos de redes estáticas

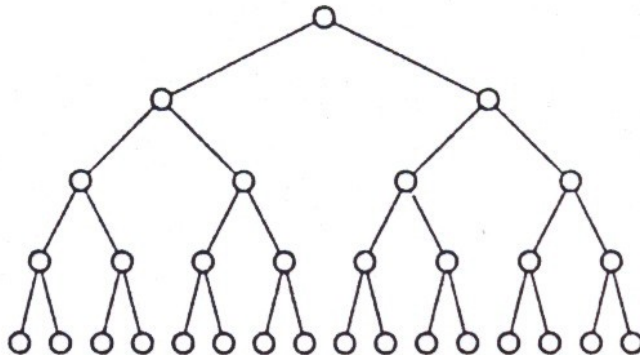


Bartel shifter



malha completa

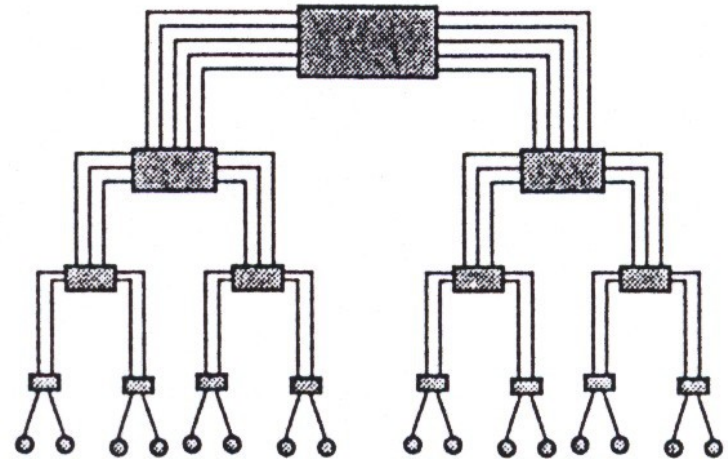
Exemplos de redes estáticas



(a) Binary tree

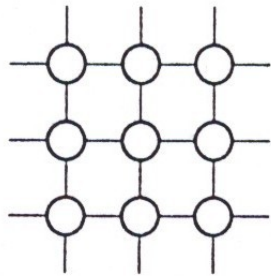


(b) Star

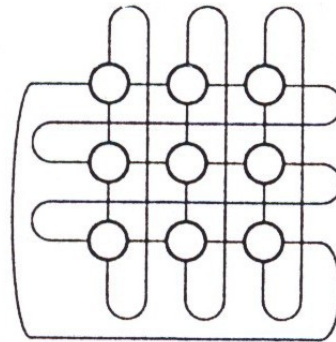


(c) Binary fat tree

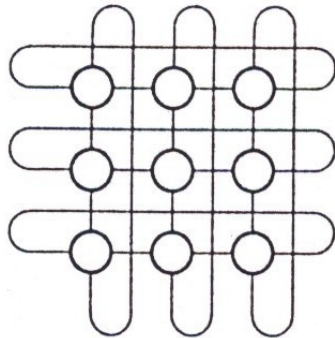
Exemplos de redes estáticas



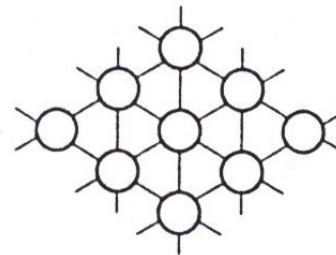
(a) Mesh



(b) Illiac mesh

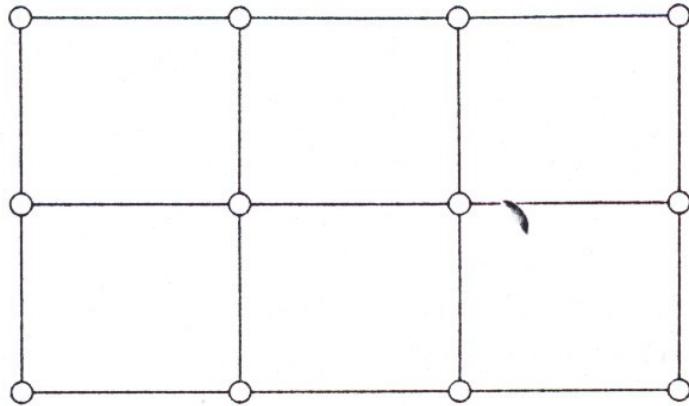


(c) Torus

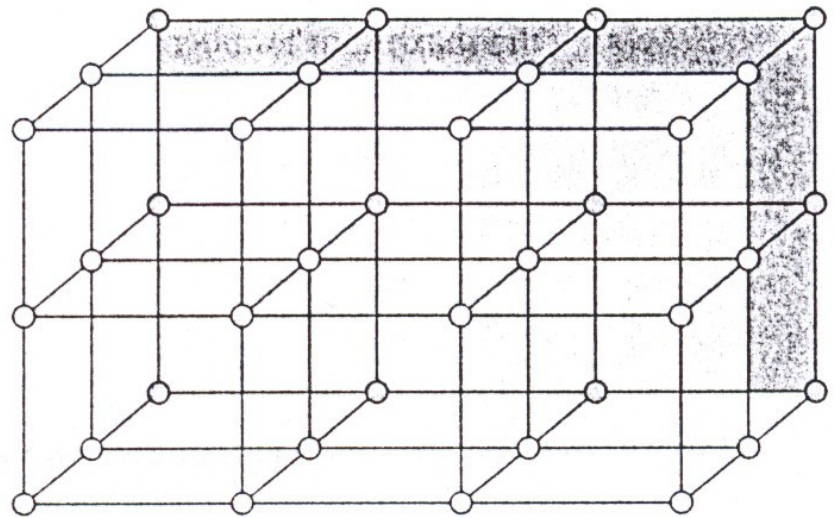


(d) Systolic array

Exemplos de redes estáticas



(a)

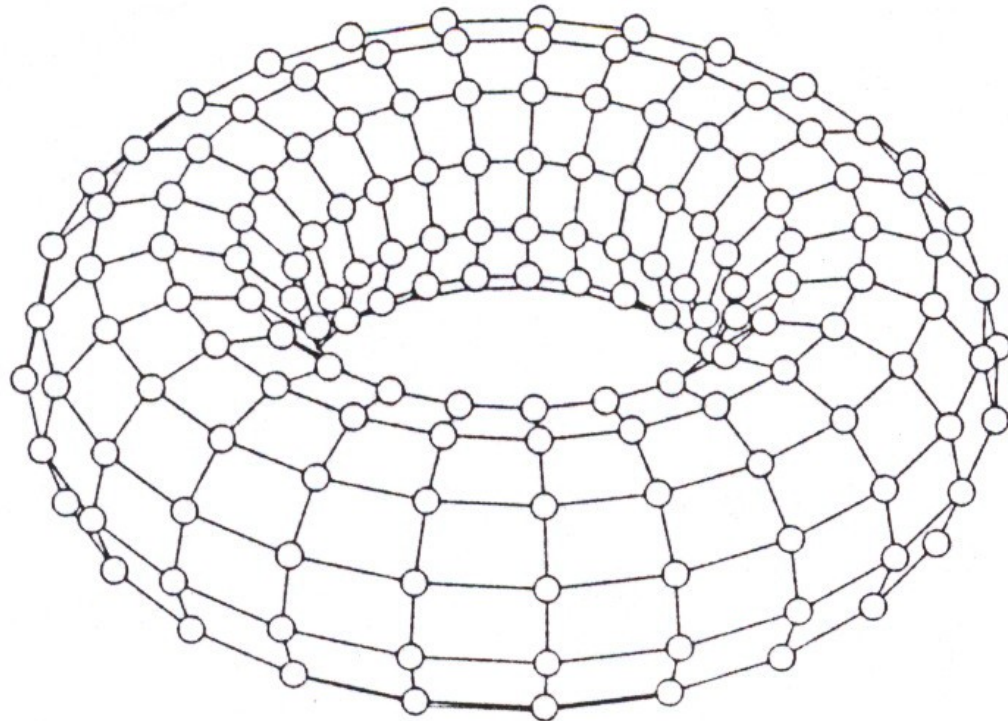


(b)

Exemplos de redes estáticas

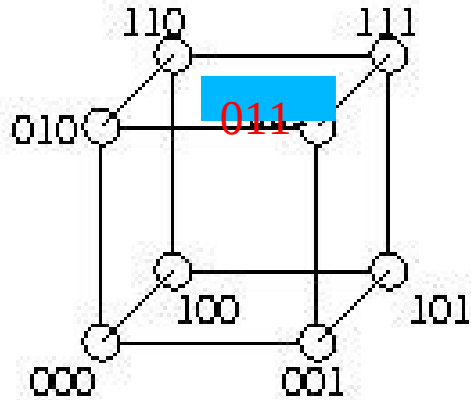


■ Toróide

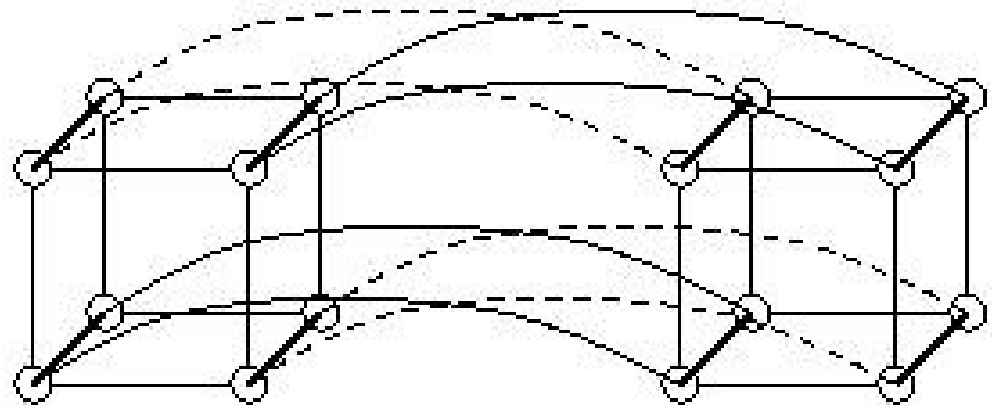


(c)

Exemplos de redes estáticas



Hipercubo de grau 3

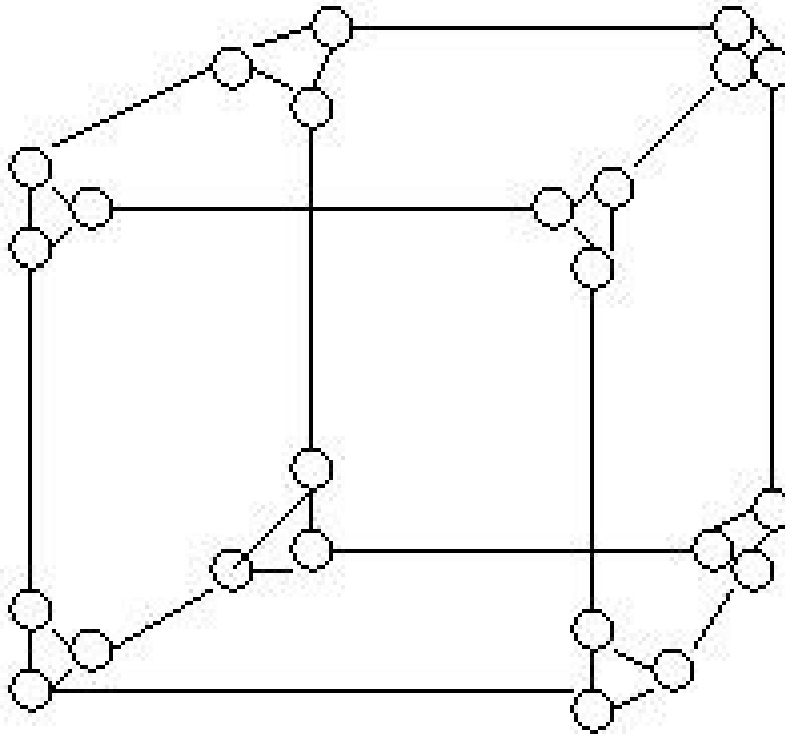


Hipercubo de grau 4

Exemplos de redes estáticas



|



CUBO 3-Connecto

|

Sumário de características de redes estáticas



Table 2.2 Summary of Static Network Characteristics

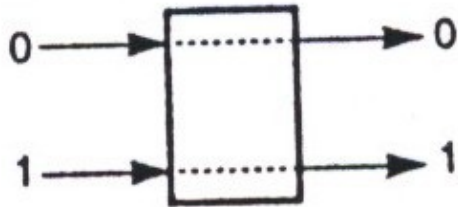
Network type	Node degree, d	Network diameter, D	No. of links, l	Bisection width, B	Symmetry	Remarks on network size
Linear Array	2	$N - 1$	$N - 1$	1	No	N nodes
Ring	2	$\lfloor N/2 \rfloor$	N	2	Yes	N nodes
Completely Connected	$N - 1$	1	$N(N - 1)/2$	$(N/2)^2$	Yes	N nodes
Binary Tree	3	$2(h - 1)$	$N - 1$	1	No	Tree height $h = \lceil \log_2 N \rceil$
Star	$N - 1$	2	$N - 1$	$\lfloor N/2 \rfloor$	No	N nodes
2D-Mesh	4	$2(r - 1)$	$2N - 2r$	r	No	$r \times r$ mesh where $r = \sqrt{N}$
Illiac Mesh	4	$r - 1$	$2N$	$2r$	No	Equivalent to a chordal ring of $r = \sqrt{N}$
2D-Torus	4	$2\lceil r/2 \rceil$	$2N$	$2r$	Yes	$r \times r$ torus where $r = \sqrt{N}$
Hypercube	n	n	$nN/2$	$N/2$	Yes	N nodes, $n = \log_2 N$ (dimension)
CCC	3	$2k - 1 + \lfloor k/2 \rfloor$	$3N/2$	$N/(2k)$	Yes	$N = k \times 2^k$ nodes with a cycle length $k \geq 3$
k -ary n -cube	$2n$	$n\lfloor k/2 \rfloor$	nN	$2k^{n-1}$	Yes	$N = k^n$ nodes

Exemplos de redes dinâmicas

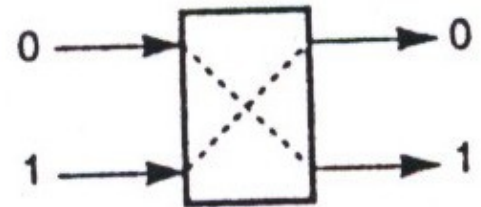


- Redes multiestágios (ômega)
- Redes Crossbar
- Redes em barramento

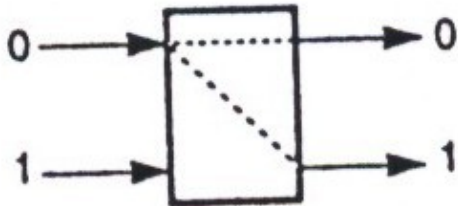
Exemplos de redes dinâmicas



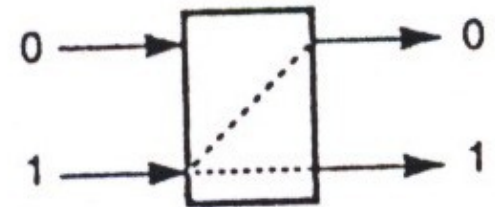
(a) Straight



(b) Crossover

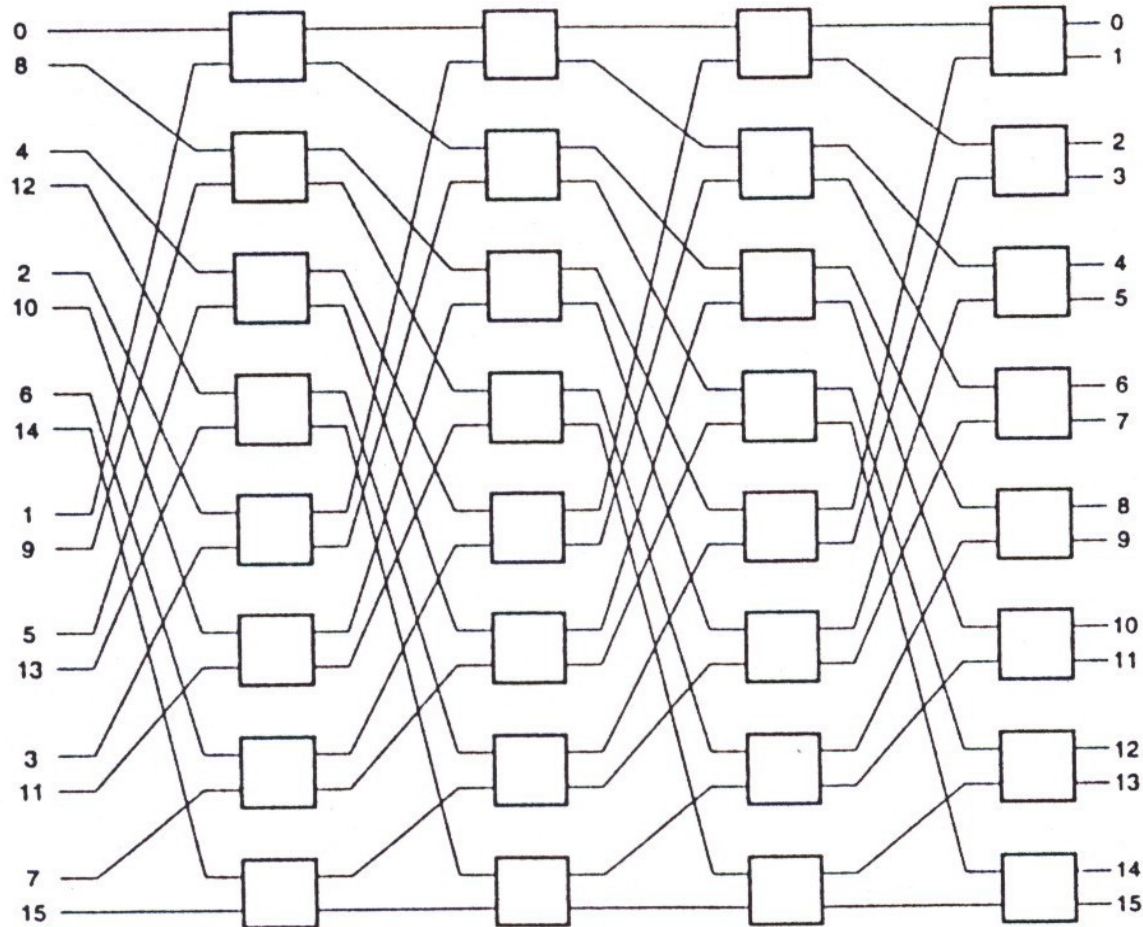


(c) Upper broadcast



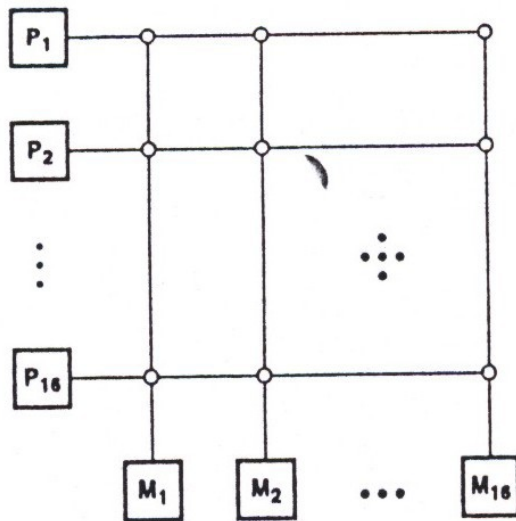
(d) Lower broadcast

Exemplos de redes dinâmicas

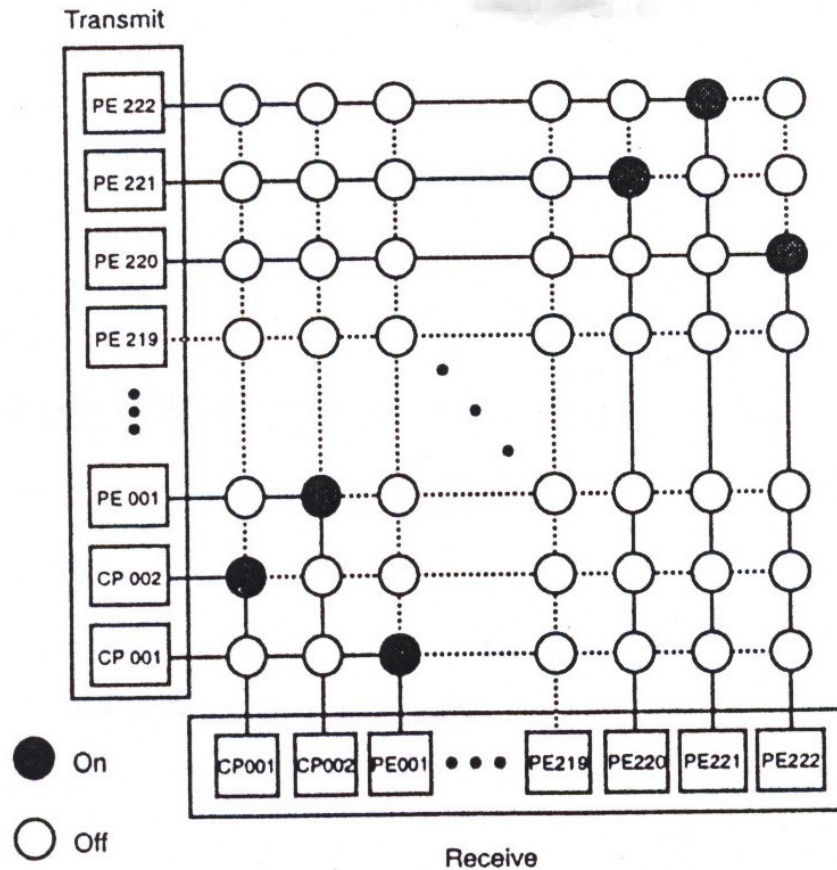


(e) A 16×16 Omega network

Exemplos de redes dinâmicas

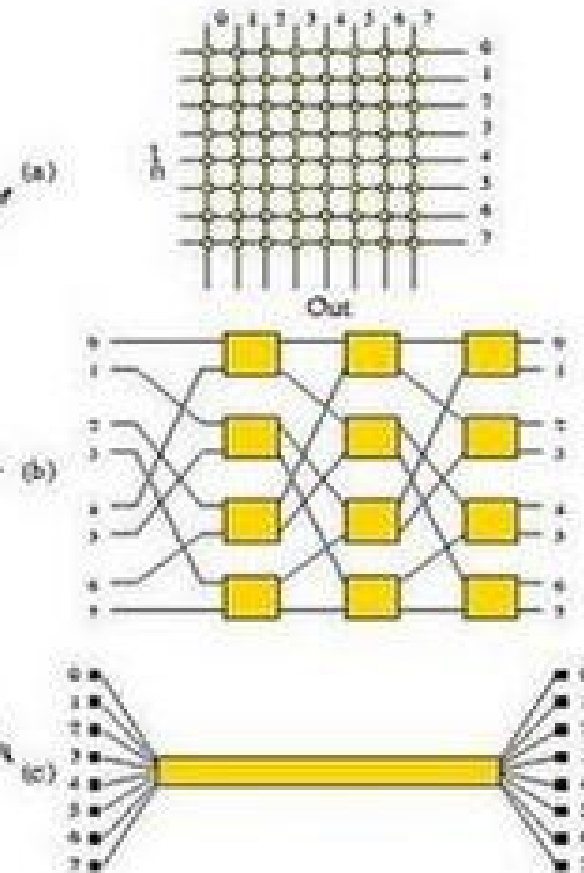
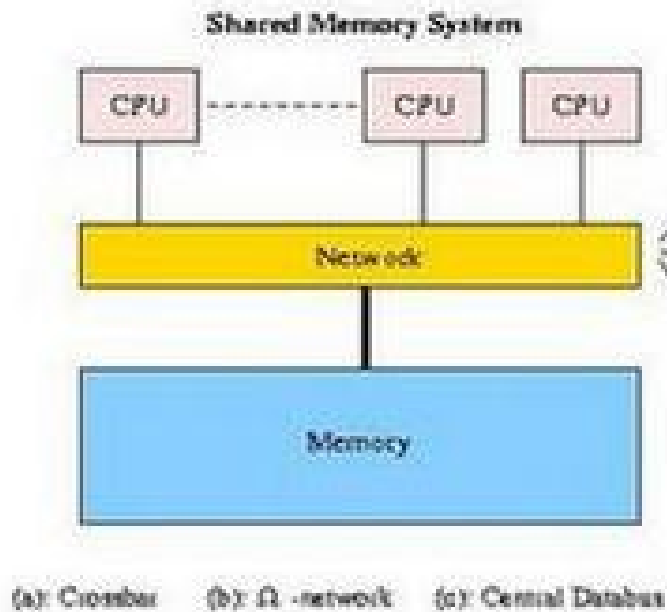


(a) Interprocessor-memory crossbar network built in the C.mmp multiprocessor at Carnegie-Mellon University (1972)



(b) The interprocessor crossbar network built in the Fujitsu VPP500 vector parallel processor (1992)

Exemplos de redes dinâmicas



Sumário de características de redes dinâmicas



Table 2.4 Summary of Dynamic Network Characteristics

Network Characteristics	Bus System	Multistage Network	Crossbar Switch
Minimum latency for unit data transfer	Constant	$O(\log_k n)$	Constant
Bandwidth per processor	$O(w/n)$ to $O(w)$	$O(w)$ to $O(nw)$	$O(w)$ to $O(nw)$
Wiring Complexity	$O(w)$	$O(nw \log_k n)$	$O(n^2 w)$
Switching Complexity	$O(n)$	$O(n \log_k n)$	$O(n^2)$
Connectivity and routing capability	Only one to one at a time.	Some permutations and broadcast, if network unblocked	All permutations, one at a time.
Representative computers	Symmetry S-1, Encore Multimax	BBN TC-2000, IBM RP3	Cray Y-MP/816, Fujitsu VPP500
Remarks	Assume n processors on the bus; bus width is w bits.	$n \times n$ MIN using $k \times k$ switches with line width of w bits.	Assume $n \times n$ crossbar with line width of w bits.

Novos modelos de rede



■ Infiniband

- Introduzida pela Intel e um consórcio envolvendo gigantes como IBM, HP, Cisco, Microsoft e outros, a partir de 2001
- Estrutura totalmente configurável, porém a mais usada é também a *fat-tree*
- *Largura de banda de até 20 Gbits/s*

Modo de operação InfiniBand



- 16 canais lógicos (linhas virtuais) por link físico
- Uma linha para gerenciamento, restante para dados
- Dados em fluxos de pacotes
- Linha virtual dedicada temporariamente para transferências ponto-a-ponto
- Switch mapeia tráfego de linhas de entrada para linhas de saída

Novos modelos de rede



NCSA MPICH-VMI Pallas Bidir. BW

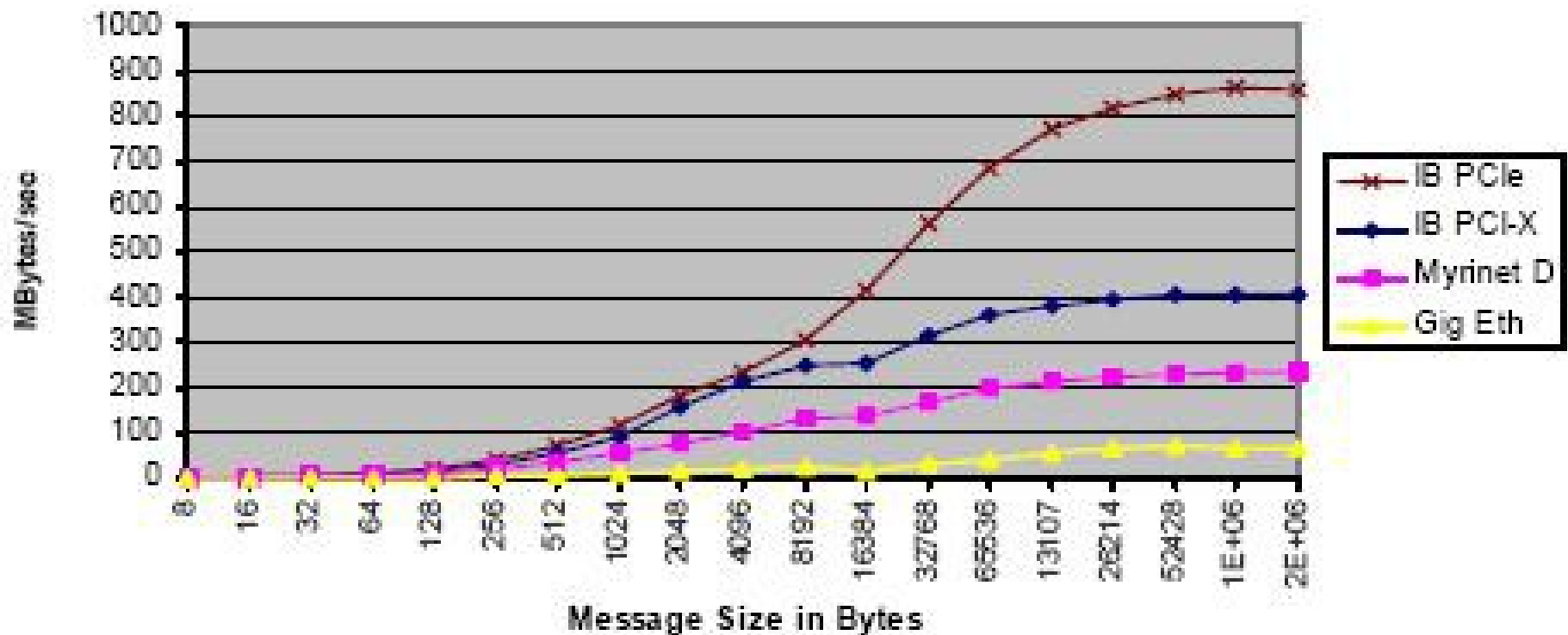


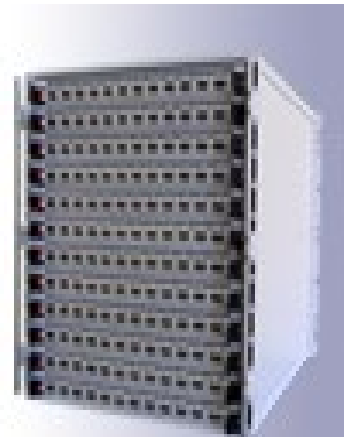
Figure 9: Aggregate 1.74GBytes/sec (870+870MB) Bandwidth

Novos modelos de rede



144 Port Switch

Modular 144-Ports
at 10Gb/sec



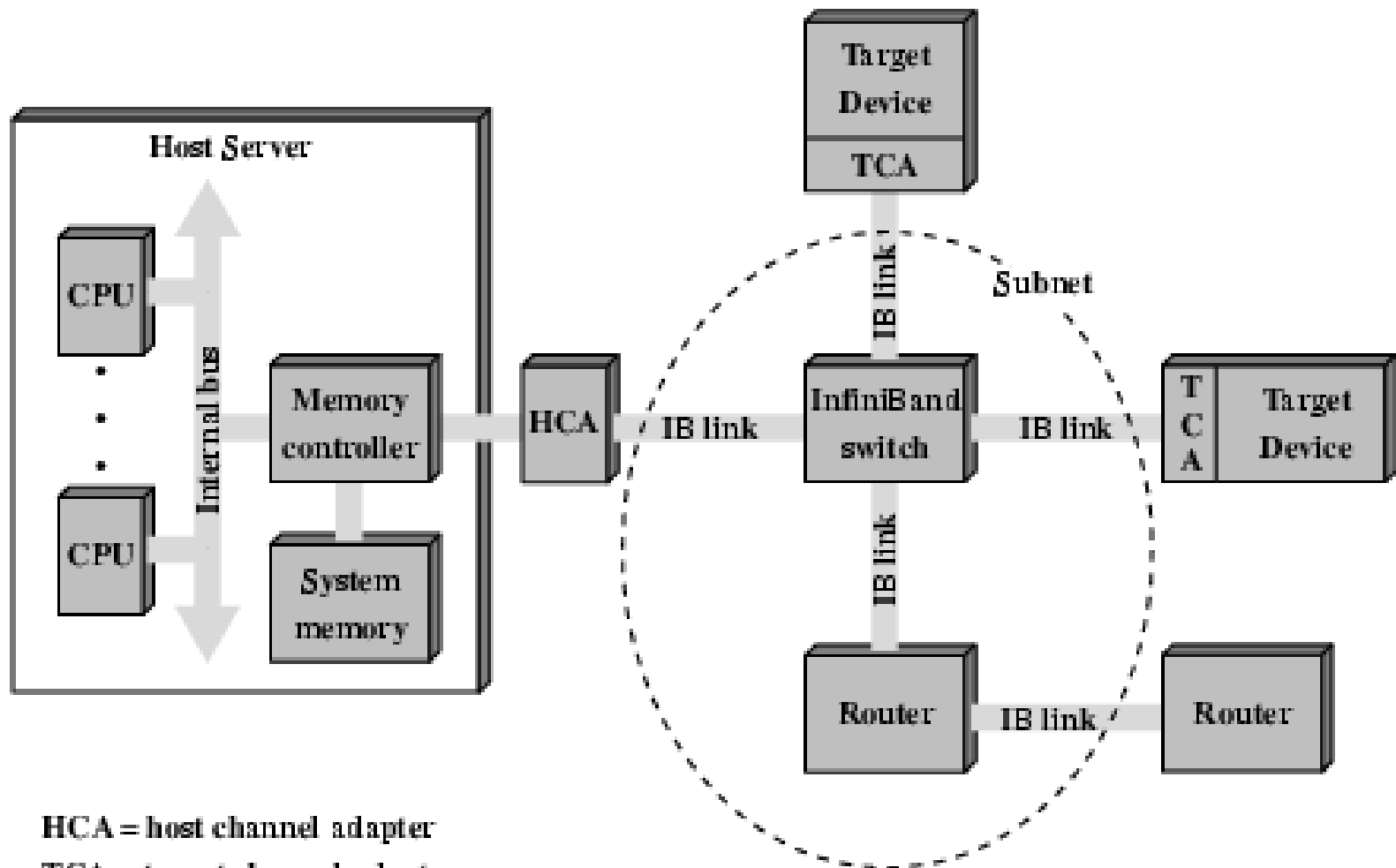
Novos modelos de rede



- Infinibanc



InfiniBand Switch Fabric



HCA = host channel adapter

TCA = target channel adapter

Novos modelos de rede



■ Infiniband

“10TFLOP” Sized Cluster



Software providers

- Fluent
- Mellanox
- OpenIB.org
- Ohio State Univ
- SDSC at UCSD
- Voltaire

Featured Technologies

- EM84T
- PCI Express*
- DDR2 Memory
- 800 MHz FSB
- Rocks EM84T CM8
- InfiniBand* Architecture
- Fluent CFD Solution

IEEE 1394 FireWire



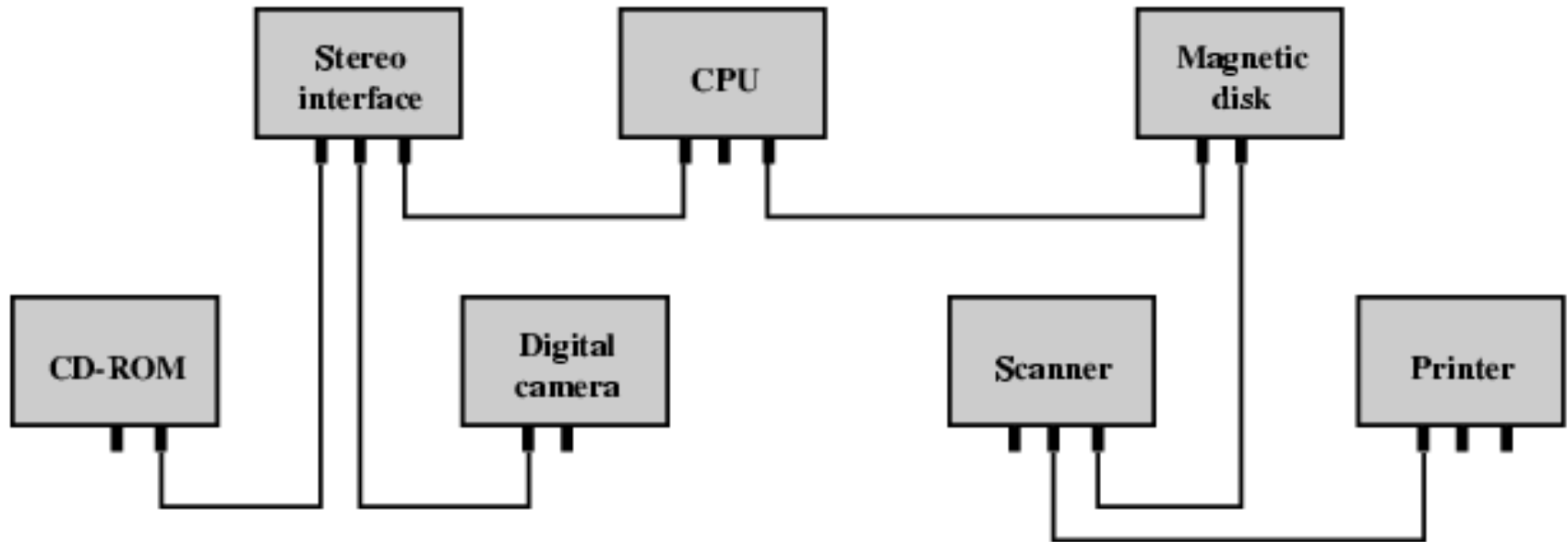
- High performance serial bus
- Fast
- Low cost
- Easy to implement
- Also being used in digital cameras, VCRs and TV

FireWire Configuration



- Daisy chain
- Up to 63 devices on single port
 - Really 64 of which one is the interface itself
- Up to 1022 buses can be connected with bridges
- Automatic configuration
- No bus terminators
- May be tree structure

Simple FireWire Configuration



Novos modelos de rede



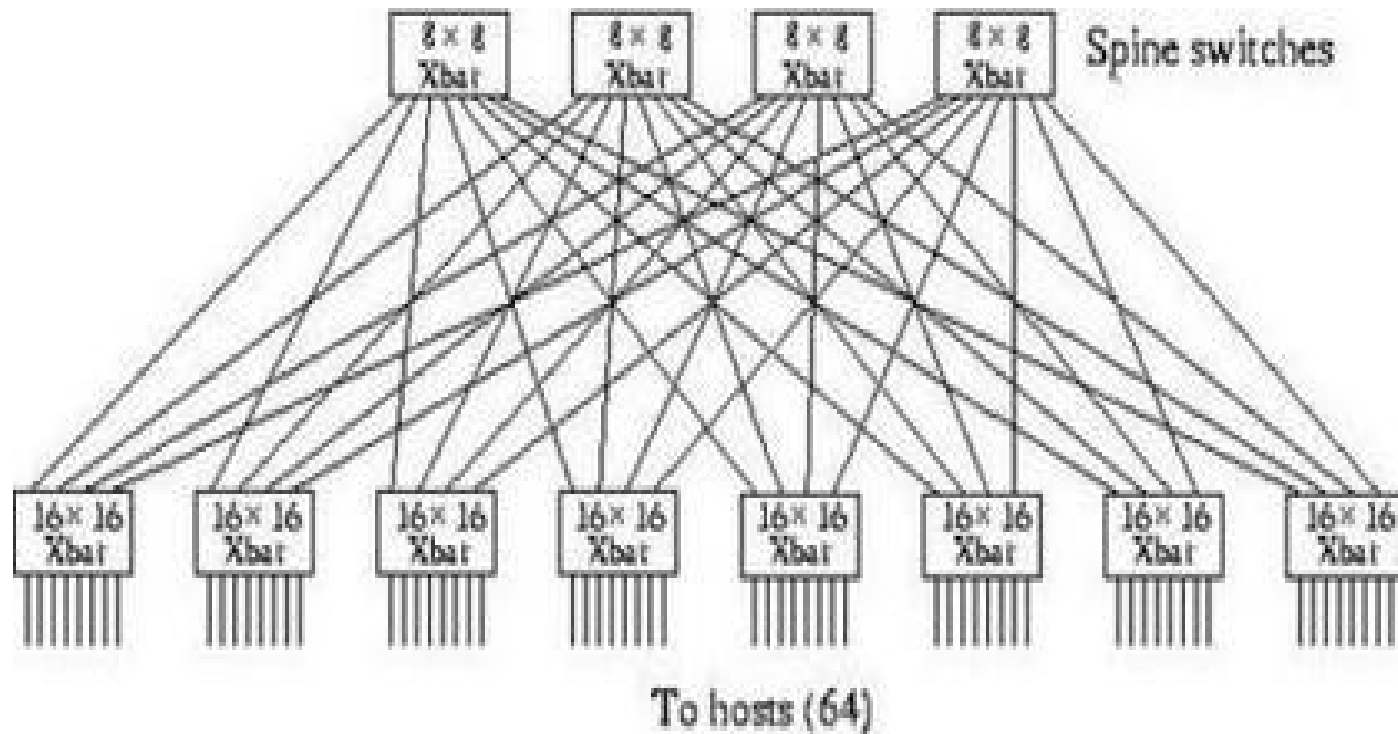
■ Myrinet

- Estrutura de rede logaritmica (Clos) em *fat-tree*
- Faz uso de acesso remoto direto a memória (RDMA), evitando overhead do SO
- Latência para msg pequenas de $2.2\mu s$ na versão MX-10G

Novos modelos de rede



■ Myrinet



Novos modelos de rede



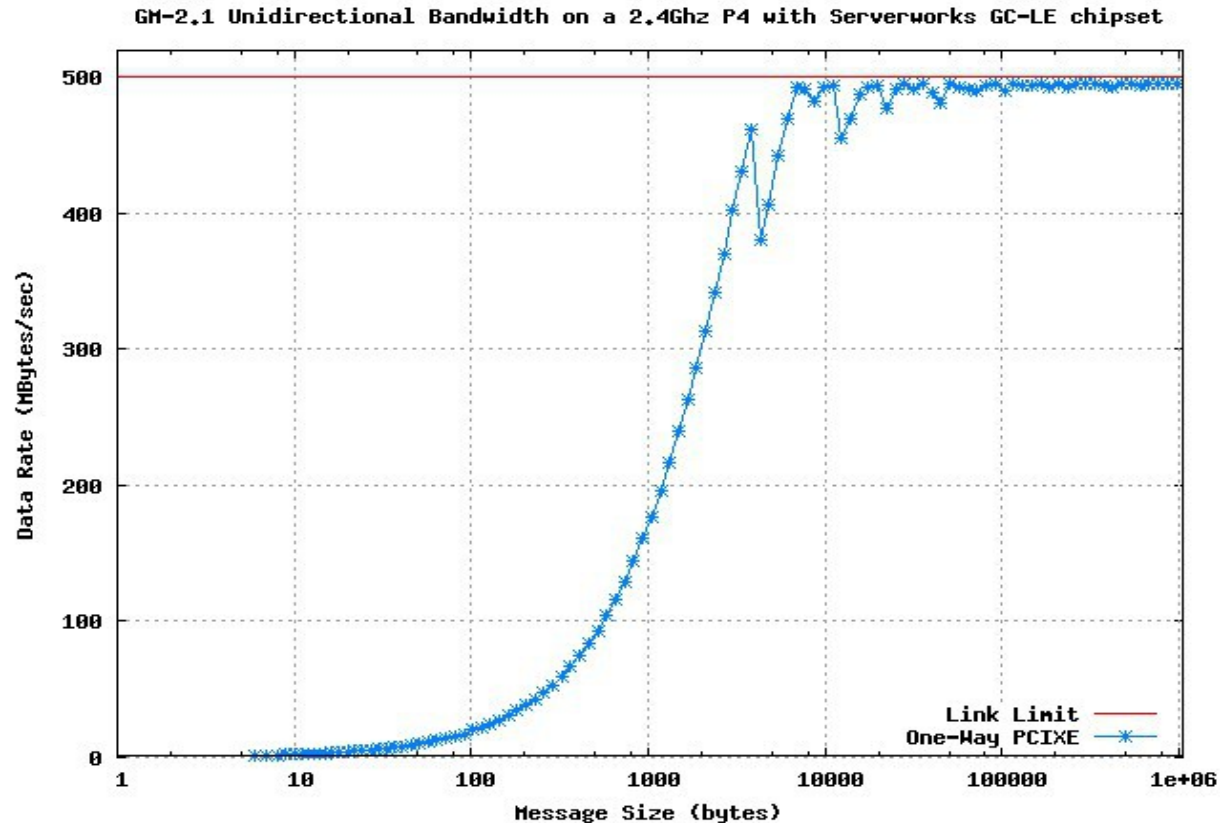
■ Myrinet



Novos modelos de rede



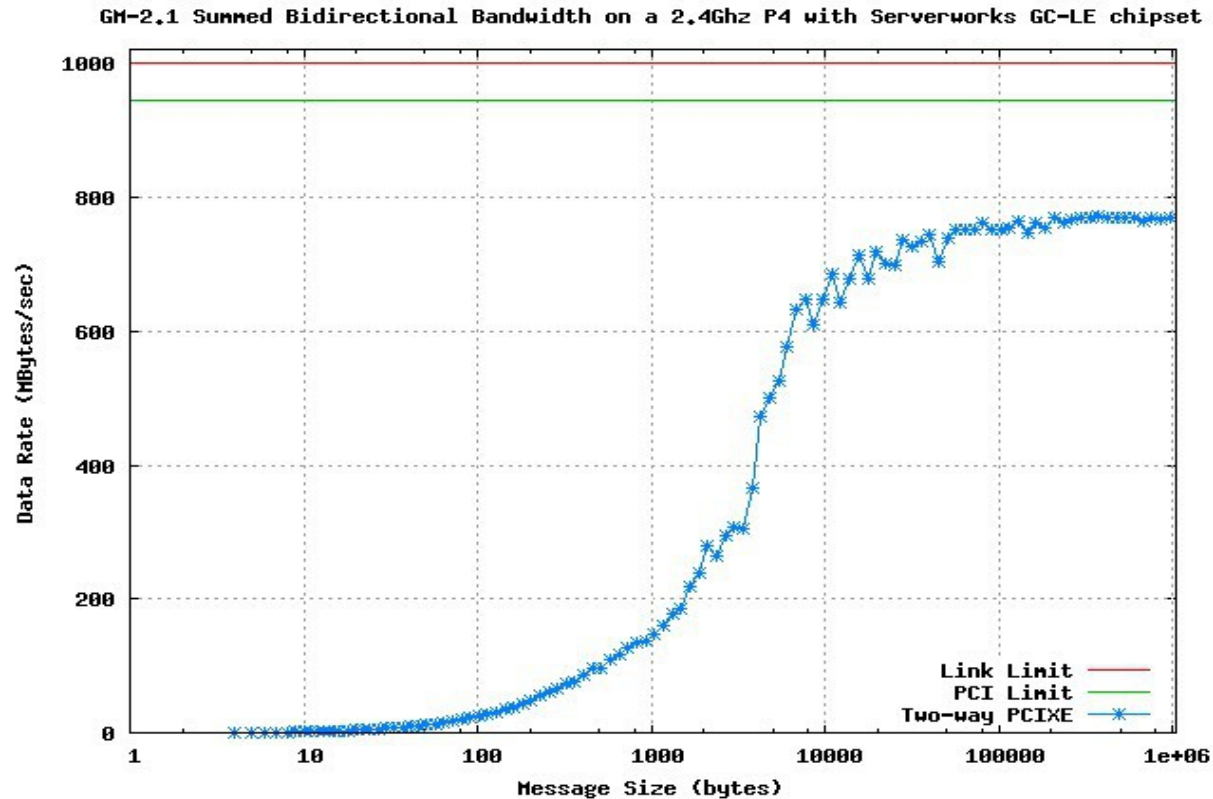
■ Desempenho unidirecional da myrinet (1G)



Novos modelos de rede



■ Desempenho bidirecional da myrinet (1G)



Novos modelos de rede



■ Custos da Myrinet 2000

- Switch 16 portas – 6.600 dolares
- Switch 256x256 – 128.000 dolares
- Placa de rede – 500 a 800 dolares
- Cabos – 75 a 200 dolares por 3 metros

Novos modelos de rede



■ QsTenG

- ☐ Produzida pela Quadrics
- ☐ Banda de passagem de 1.3 Gbyte/s
- ☐ Latência inferior a $3\mu\text{s}$
- ☐ Estrutura de *fat-tree*

Novos modelos de rede



■ QsNet



Novos modelos de rede



- QsTenG

