

Arquiteturas de computadores paralelos

Arquiteturas de computadores paralelos

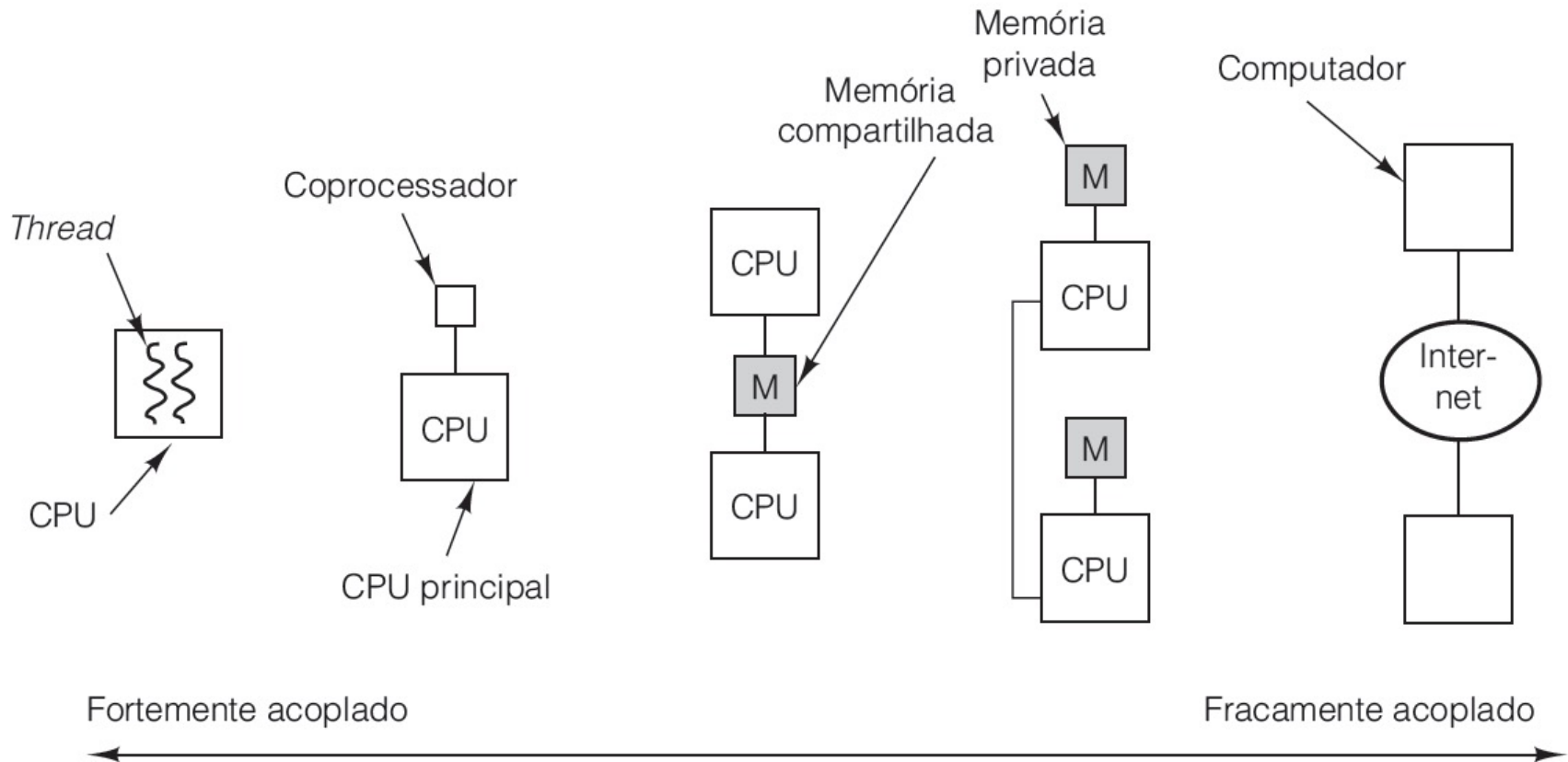
- Para enfrentar problemas cada vez maiores, os arquitetos de computadores estão recorrendo cada vez mais a computadores paralelos.
- Apesar de talvez não ser possível construir uma máquina com uma única CPU e um tempo de ciclo de 0,001 ns, pode ser perfeitamente viável produzir uma com 1.000 CPUs com um tempo de ciclo de 1 ns cada.
- Embora esse último projeto use CPUs mais lentas do que o primeiro, sua capacidade total de computação é teoricamente a mesma.

Arquiteturas de computadores paralelos

- O paralelismo pode ser introduzido em vários níveis.
- No nível mais baixo, ele pode ser adicionado ao chip da CPU, por pipeline e projetos superescalares com várias unidades funcionais.
- No nível seguinte, placas extras de CPU com capacidade de processamento adicional podem ser acrescentadas a um sistema.
- Contudo, para conseguir um fator de cem, de mil, ou de milhão, é necessário replicar CPUs inteiras e fazer que todas elas funcionem juntas com eficiência.

Arquiteturas de computadores paralelos

- Paralelismo no chip. Coprocessador. Multiprocessador. Multicomputador. Grade.



Paralelismo no chip

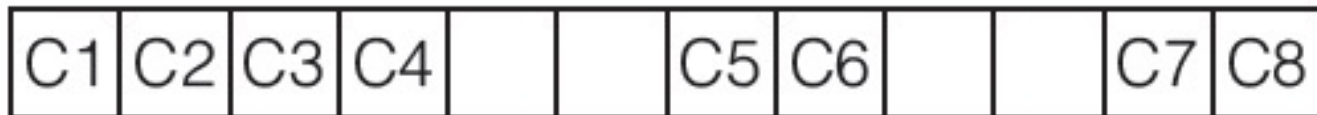
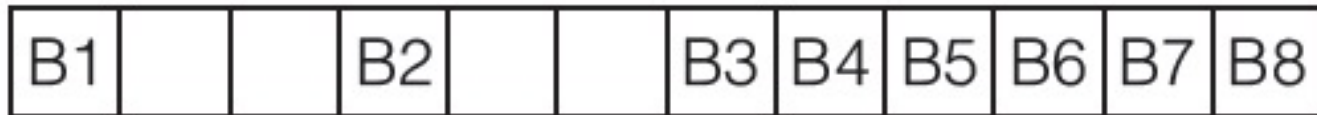
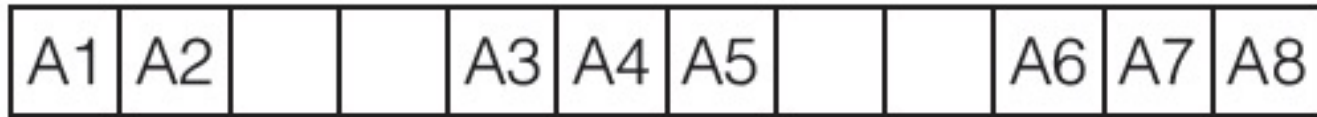
- Um modo de conseguir paralelismo no nível mais baixo é emitir múltiplas instruções por ciclo de *clock*.
- A outra forma de paralelismo no nível da instrução é encontrada em processadores VLIW.
- VLIW transfere do tempo de execução para o tempo de compilação o trabalho de determinar quais instruções podem ser emitidas em conjunto.
- Outra forma de paralelismo de baixo nível é o paralelismo no nível da memória, no qual há múltiplas operações de memória no ar ao mesmo tempo.

Paralelismo no chip

- Todas as CPUs modernas, com paralelismo (*pipeline*), têm um problema inerente:
 - quando uma referência à memória encontra uma ausência das caches de nível 1 e nível 2, há uma longa espera até que a palavra requisitada seja carregada na cache, portanto, o *pipeline* para.
- Uma abordagem para lidar com essa situação, denominada ***multithreading no chip***, permite que a CPU gerencie múltiplos *threads* de controle ao mesmo tempo em uma tentativa de mascarar essas protelações.

Paralelismo no chip

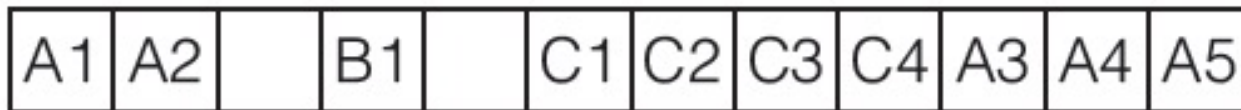
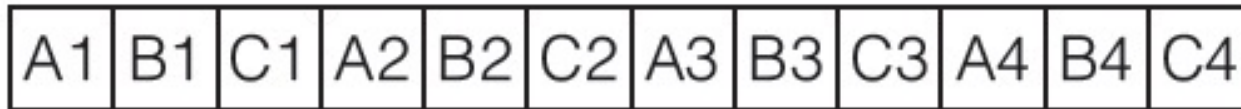
- Três *threads*.



Ciclo →

Paralelismo no chip

- *Multithreading* de granulação fina. *Multithreading* de granulação grossa.



Ciclo →

Paralelismo no chip

- *Multithreading* com uma CPU superescalar de emissão dual. *Multithreading* de granulação fina. *Multithreading* de granulação grossa. *Multithreading* simultâneo.

A1	B1	C1	A3	B2	C3	A5	B3	C5	A6	B5	C7
A2		C2	A4		C4		B4	C6	A7	B6	C8

Ciclo →

A1	B1	C1	C3	A3	A5	B2	C5	A6	A8	B3	B5
A2		C2	C4	A4			C6	A7		B4	B6

Ciclo →

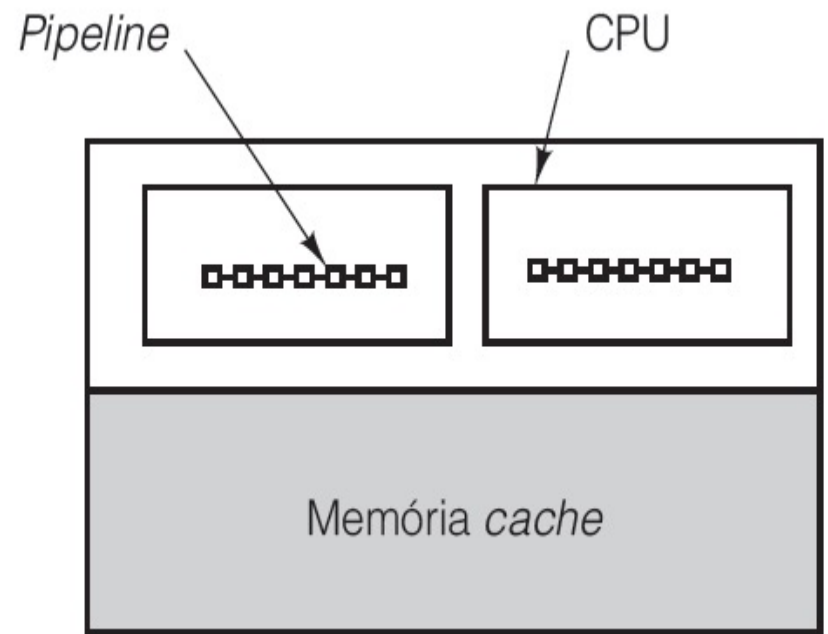
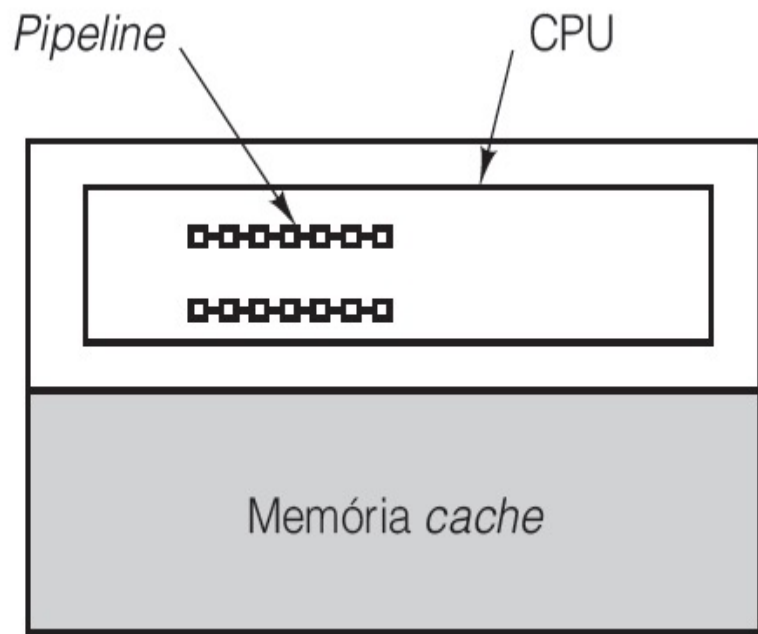
A1	B1	C2	C4	A4	B2	C6	A7	B3	B5	B7	C7
A2	C1	C3	A3	A5	C5	A6	A8	B4	B6	B8	C8

Ciclo →

Paralelismo no chip

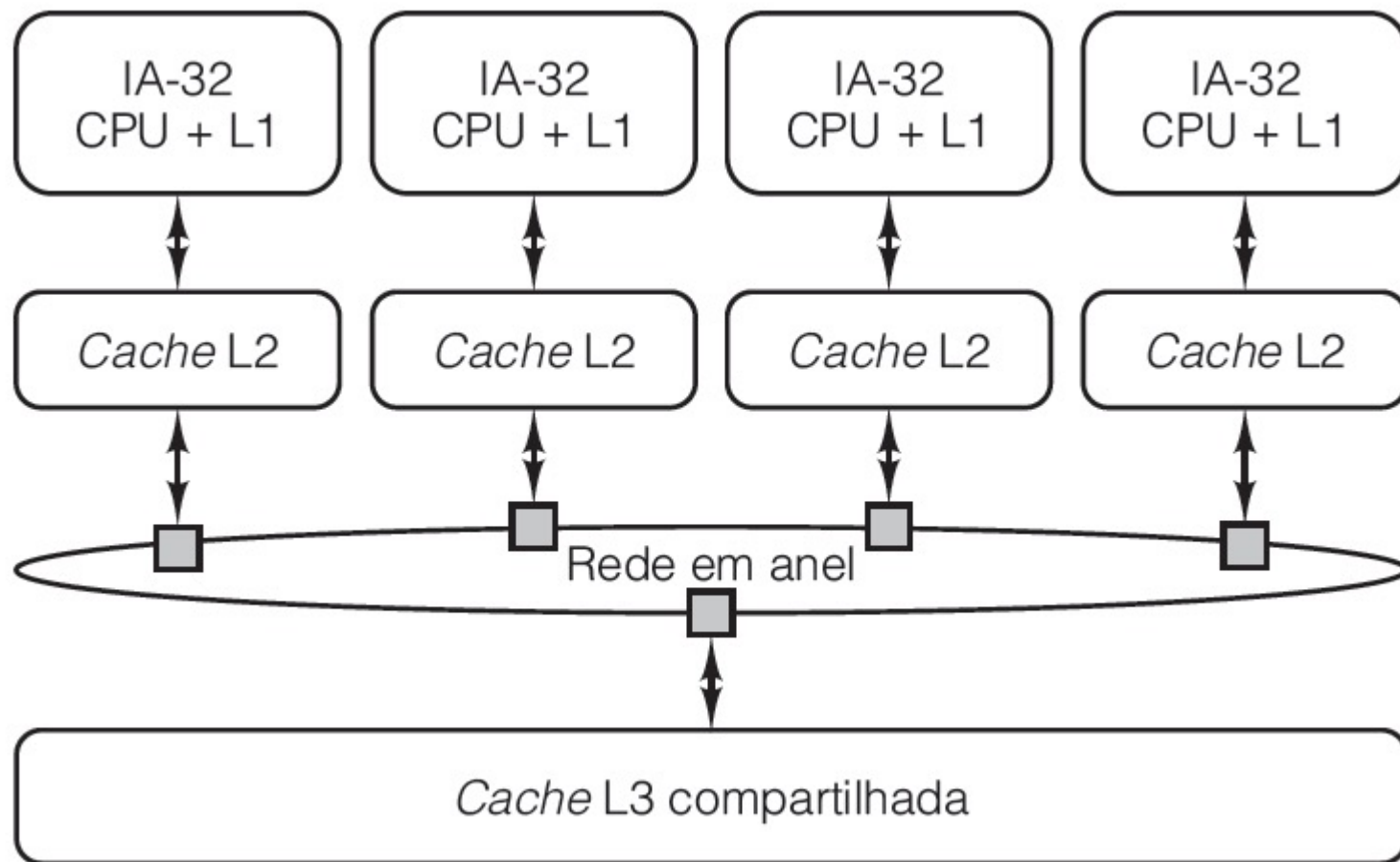
- Embora o *multithreading* ofereça ganhos em desempenho significativos por um custo modesto, para algumas aplicações é preciso um ganho em desempenho muito maior do que ele pode oferecer.
- Para conseguir esse desempenho estão sendo desenvolvidos chips multiprocessadores.
- Há dois projetos predominantes para multiprocessadores de pequena escala em um único chip:
 - Chip com pipeline dual.
 - Chip com dois núcleos.

Paralelismo no chip



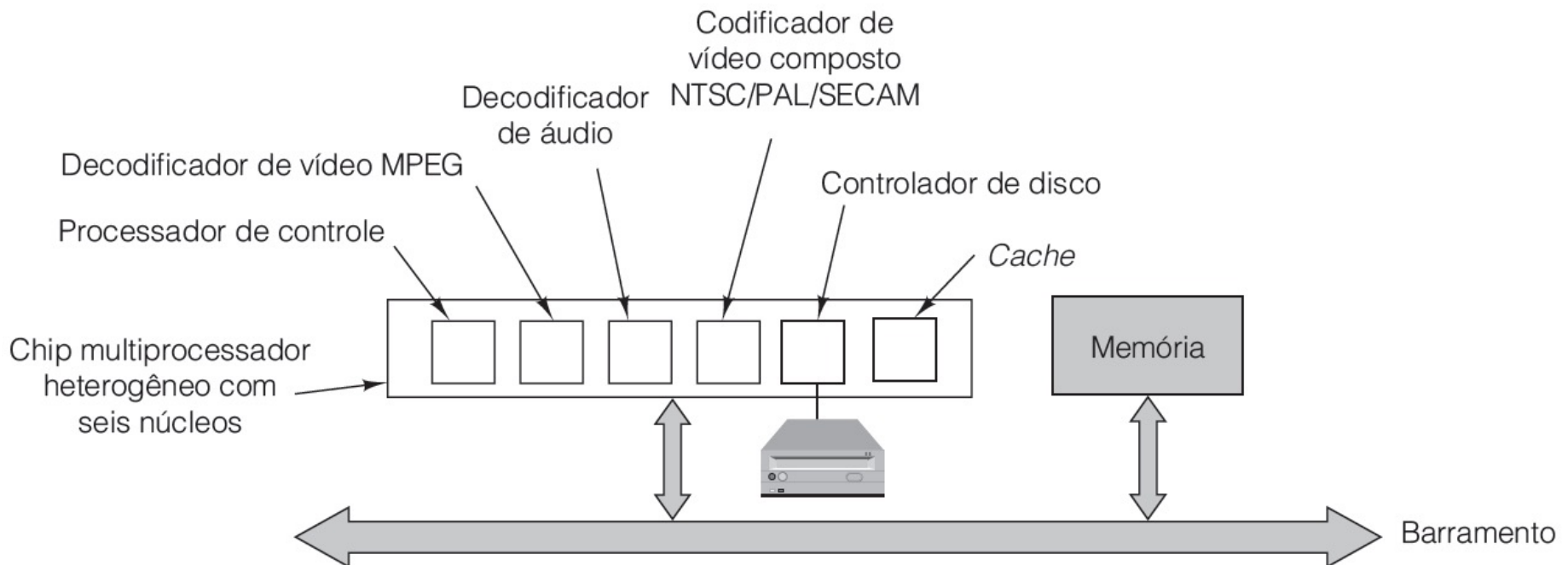
Paralelismo no chip

- Arquitetura do multiprocessador em um único chip do Core i7.



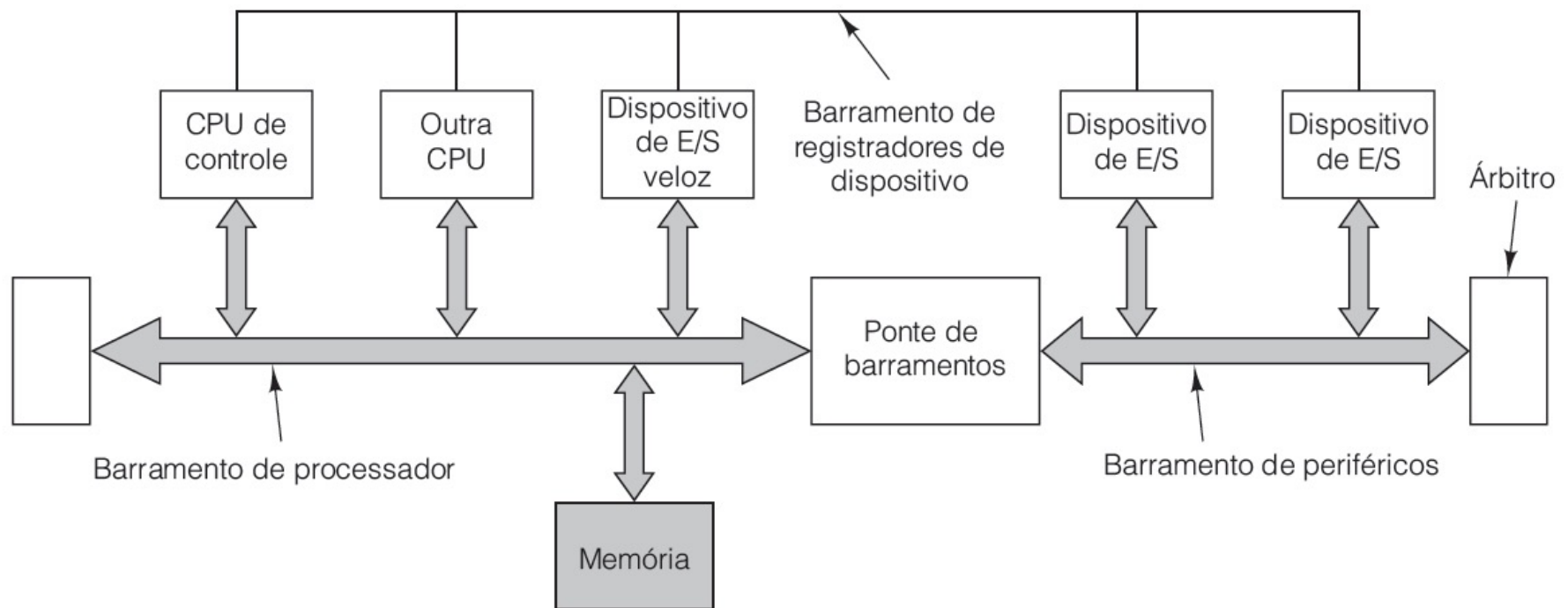
Paralelismo no chip

- A estrutura lógica de um simples aparelho de DVD contém um multiprocessador heterogêneo com múltiplos núcleos para diferentes funções.



Paralelismo no chip

- Exemplo da arquitetura CoreConnect da IBM.

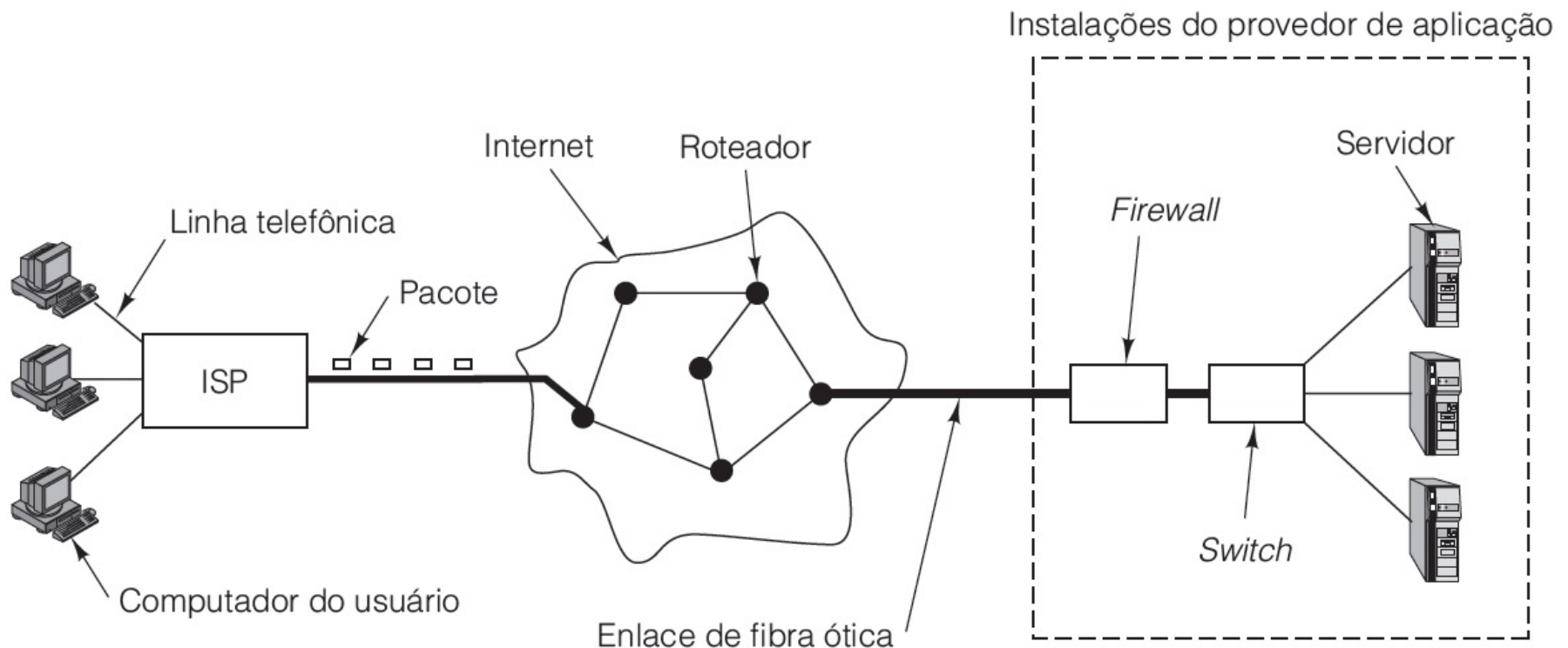


Coprocessadores

- Coprocessadores podem variar de um gabinete separado (os canais de E/S do 360) a uma placa de expansão (processadores de rede) ou uma área no chip principal (ponto flutuante).
- Em todos os casos, o que os distingue é o fato de que algum outro processador é o principal e que os coprocessadores estão lá para ajudá-lo.
- Examinaremos três áreas em que é possível aumentar a velocidade: processamento de rede, multimídia e criptografia.

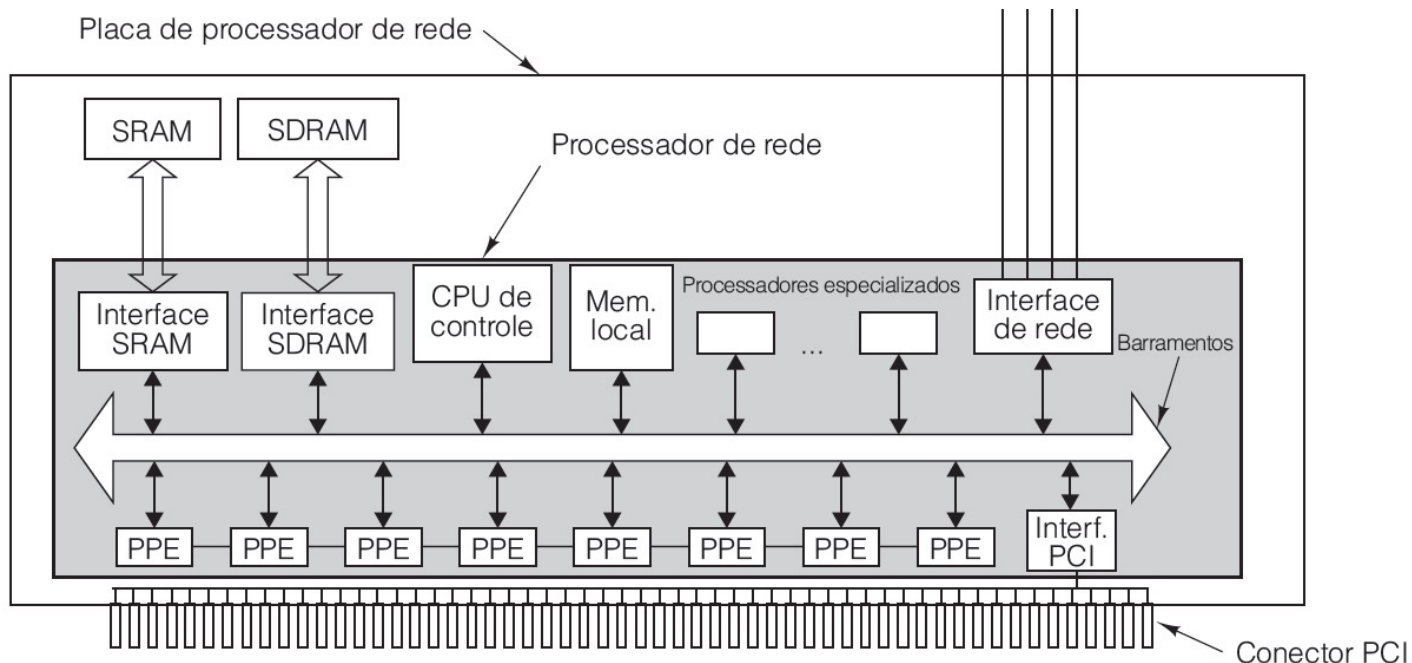
Processadores de rede

- Como os usuários são conectados a servidores na Internet.



Processadores de rede

- Processadores de rede são dispositivos programáveis que podem manipular pacotes que chegam e que saem à velocidade dos fios, isto é, em tempo real.
- Placa e chip de um processador de rede típico.



Processadores de rede

- Quando um pacote chega, ele passa por vários estágios de processamento:
 - Soma de verificação.
 - Extração do campo.
 - Classificação de pacotes.
 - Seleção de caminho.
 - Determinação da rede de destino.
 - Consulta de rota.
 - Fragmentação e reconstrução.
 - Computação.
 - Gerenciamento de cabeçalho.
 - Gerenciamento de fila.

Processadores de rede

- Geração de soma de verificação.
 - Contabilidade.
 - Coleta de dados estatísticos.
-
- Desempenho é o que importa em processadores de rede.
 - O que pode ser feito para melhorá-lo?
 - O modo mais direto de melhorar o desempenho é aumentar a velocidade de *clock* do processador de rede.
 - Introduzir mais PPEs e paralelismo costuma ser um método que dá ótimos resultados.

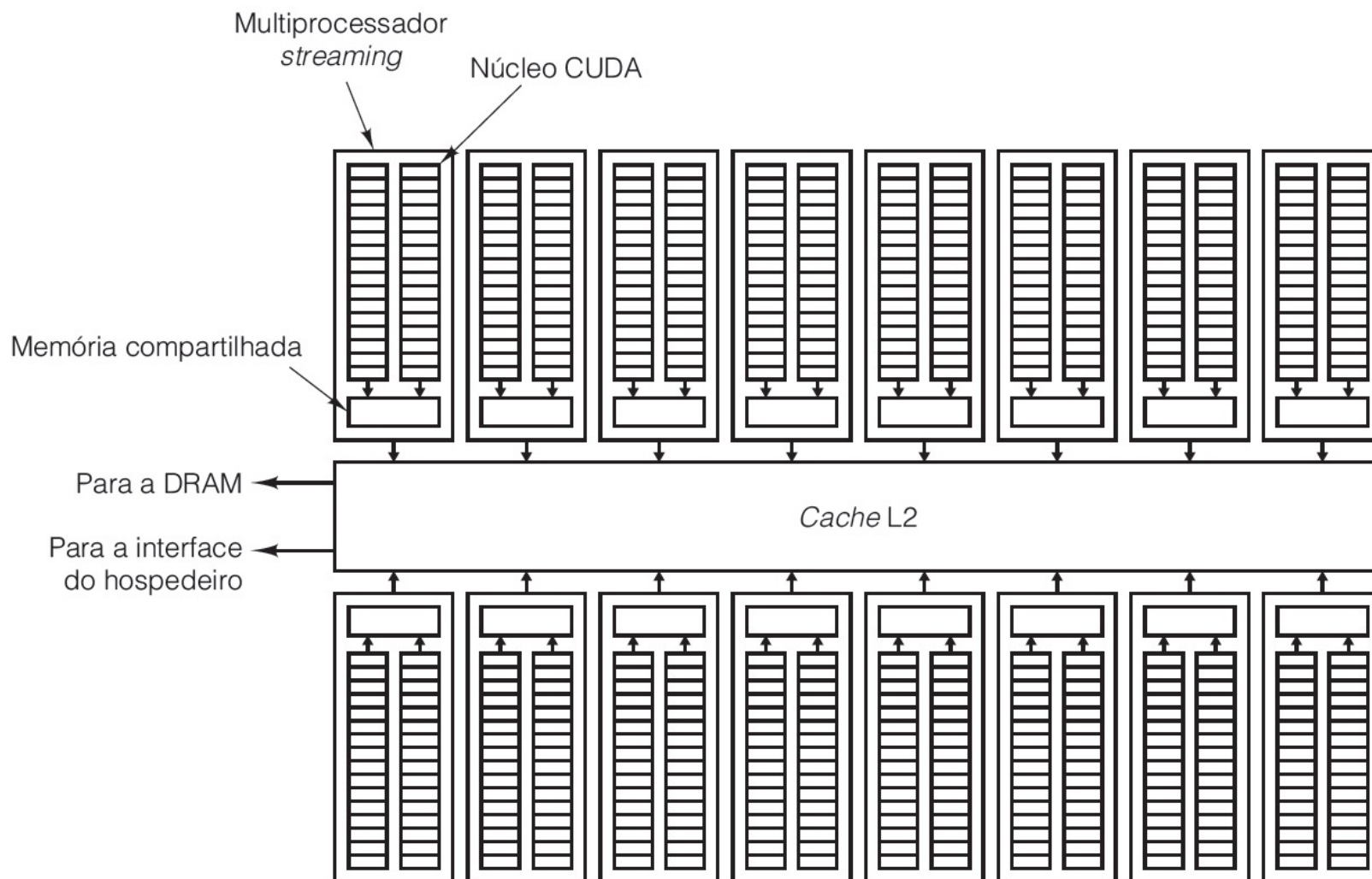
Processadores de rede

- Outra técnica é adicionar processadores especializados ou ASICs para tratar operações específicas.
- Adicionar mais barramentos internos e aumentar a largura dos barramentos existentes pode ajudar a ganhar velocidade porque os pacotes passam pelo sistema com maior rapidez.
- Por fim, substituir SDRAM por SRAM costuma ser entendido como algo que melhora o desempenho, mas, por certo, tem um preço.

Processadores gráficos

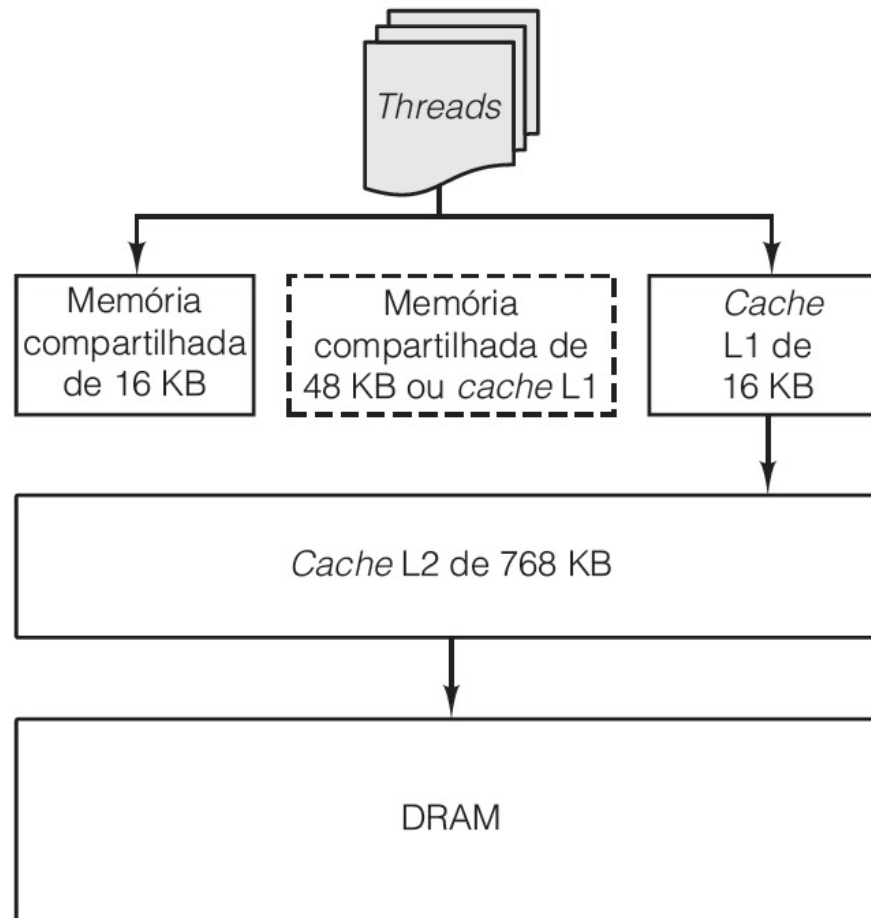
- Uma segunda área na qual coprocessadores são usados é o tratamento de processamento gráfico de alta resolução, como renderização 3D.
- Alguns PCs atuais e a maioria dos PCs futuros serão equipados com **GPUs** (*Graphics Processing Units* – unidades de processamento gráfico) para os quais passarão grandes porções do processamento geral.
- A GPU Fermi NVIDIA é uma arquitetura usada em uma família de chips de processamento gráfico que estão disponíveis em diversas velocidades e tamanhos.

Processadores gráficos



Processadores gráficos

- Hierarquia de memória da GPU Fermi.

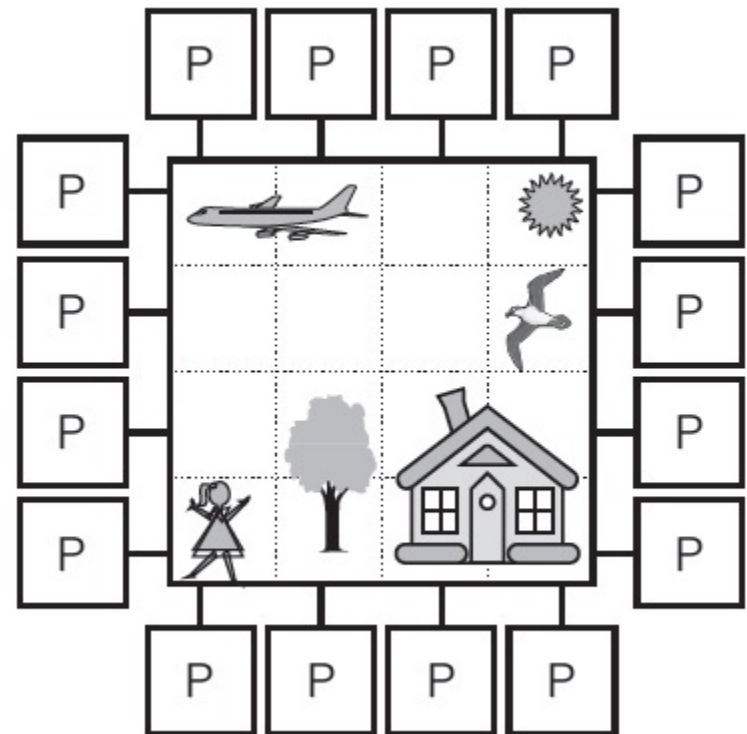
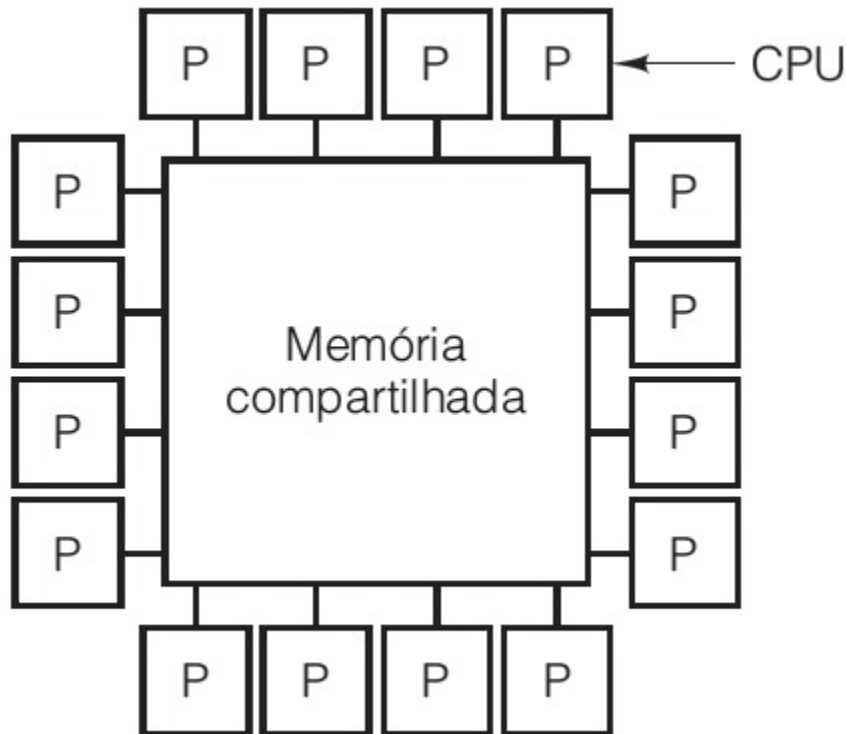


Criptoprocessadores

- Uma terceira área na qual os coprocessadores são populares é segurança, em especial segurança em redes.
- Há dois tipos gerais de criptografia:
- Criptografia de chave simétrica - é baseada na mistura completa de bits, algo equivalente a jogar uma mensagem dentro de um liquidificador.
- Criptografia de chave pública - é baseada em multiplicação e exponenciação de grandes números (por exemplo, 1.024 bits) e consome enormes quantidades de tempo.

Multiprocessadores *versus* multicomputadores

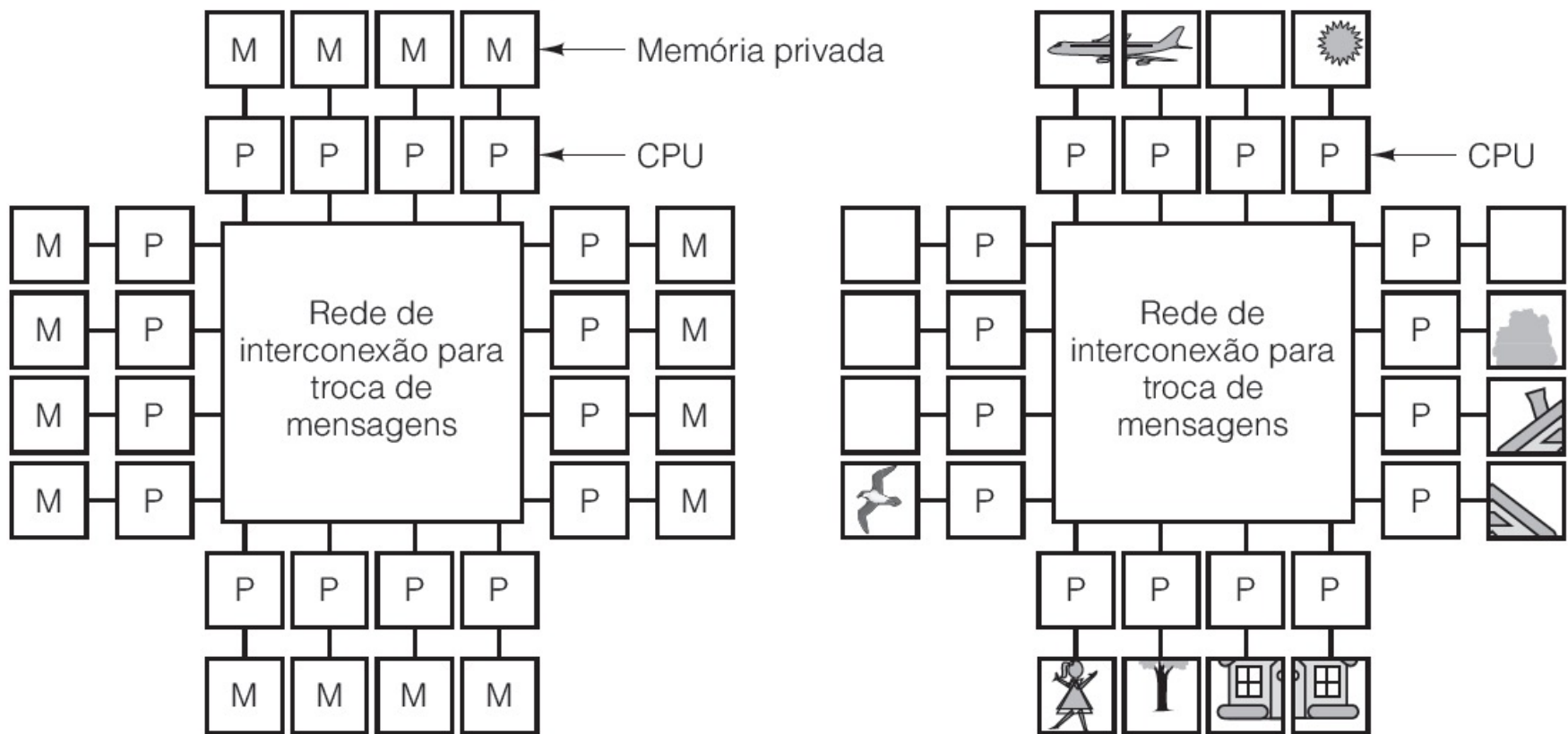
- Um computador paralelo no qual todas as CPUs compartilham uma memória comum é denominado um multiprocessador.



Multiprocessadores *versus* multicomputadores

- O aspecto fundamental de um multicomputador que o distingue de um multiprocessador é que a CPU em um multicomputador tem sua própria memória local privada, a qual pode acessar apenas executando instruções LOAD e STORE, mas que nenhuma outra CPU pode acessar usando instruções LOAD e STORE.
- Assim, multiprocessadores têm um único espaço de endereço físico compartilhado por todas as CPUs, ao passo que multicomputadores têm um espaço de endereço físico para cada CPU.

Multiprocessadores *versus* multicomputadores



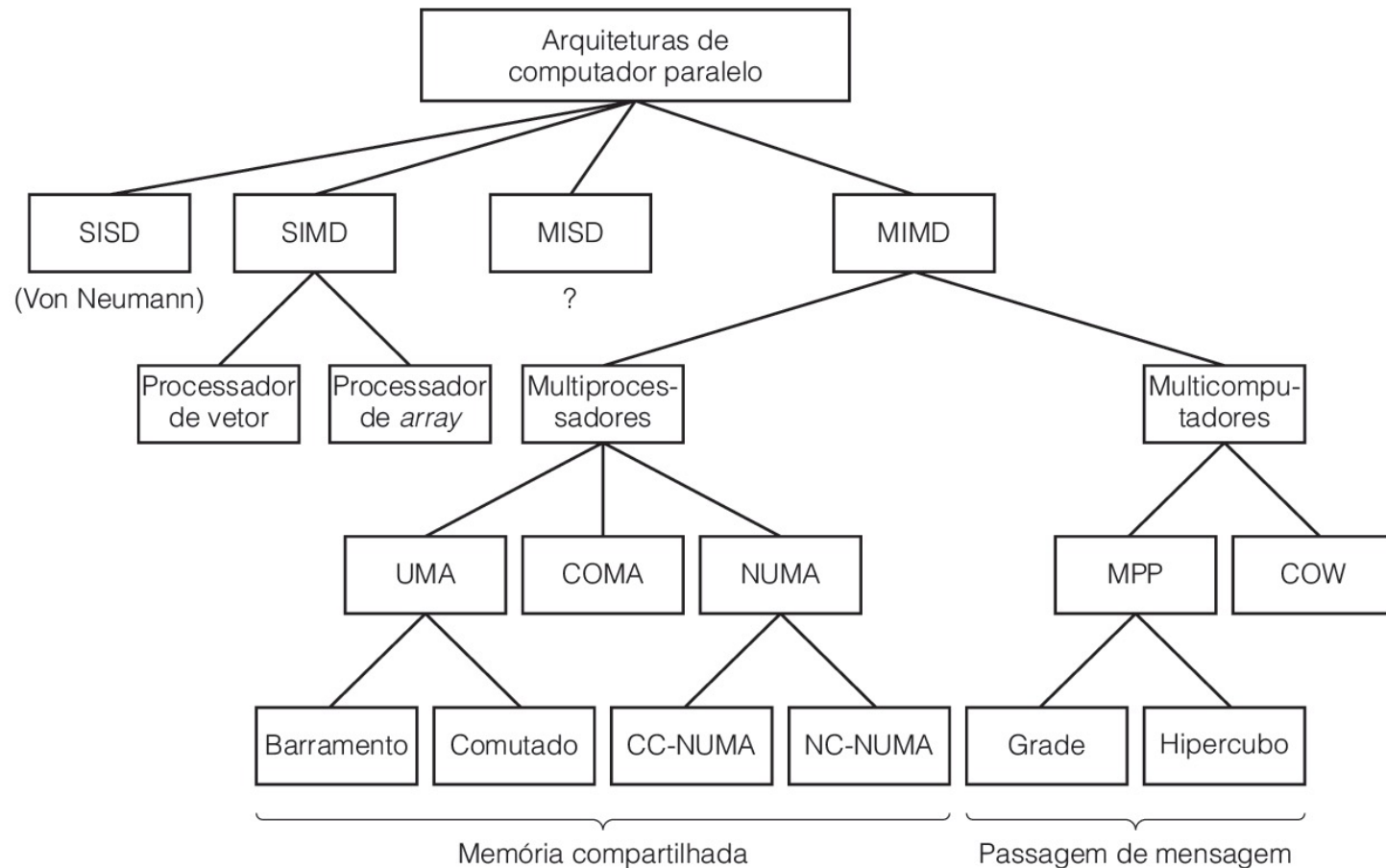
Taxonomia de computadores paralelos

- Taxonomia de Flynn para computadores paralelos.

Fluxos de instruções	Fluxos de dados	Nome	Exemplos
1	1	SISD	Máquina clássica de Von Neumann
1	Múltiplos	SIMD	Supercomputador vetorial, processador de <i>array</i>
Múltiplos	1	MISD	Possivelmente nenhum
Múltiplos	Múltiplos	MIMD	Multiprocessador, multicomputador

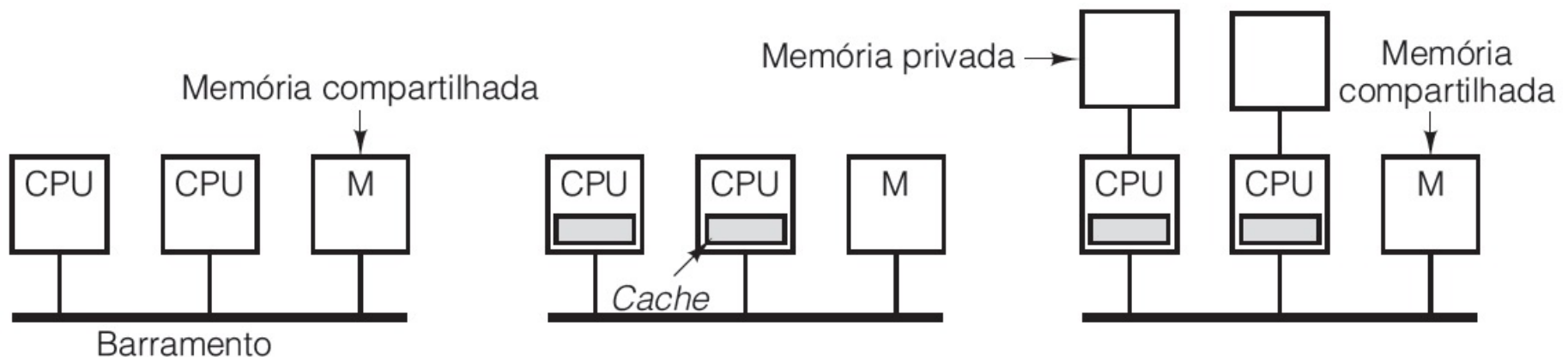
Taxonomia de computadores paralelos

- Taxonomia de computadores paralelos.



Arquiteturas de multiprocessadores simétricos UMA

- Três multiprocessadores baseados em barramento.
 - Sem *cache*.
 - Com *cache*.
 - Com *cache* e memórias privadas.



Caches de escuta

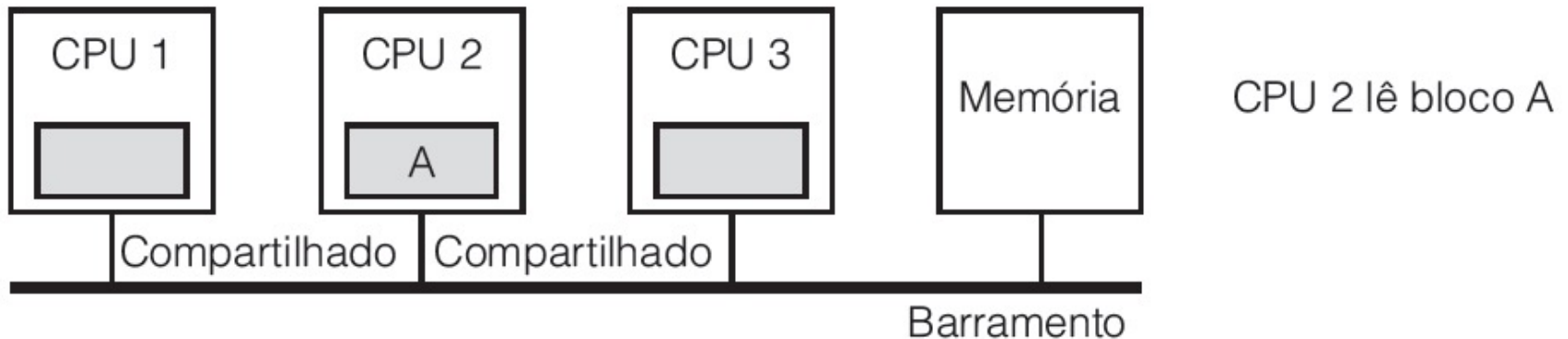
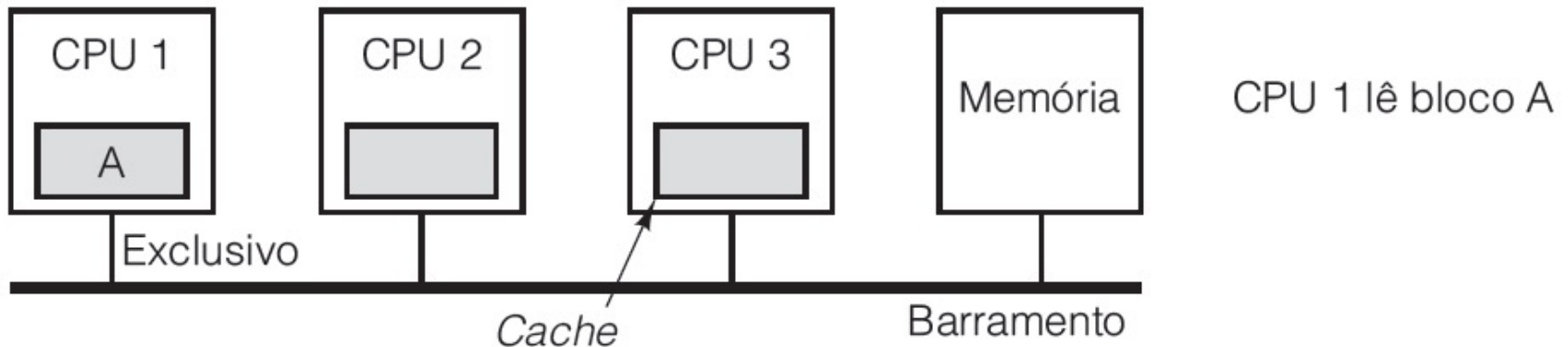
- O controlador de *cache* é projetado especialmente para permitir que ele escute o barramento monitorando todas as requisições de barramento de outras CPUs e *caches* e execute alguma ação em certos casos.
- Esses dispositivos são denominados ***caches de escuta*** ou ***caches de espia*** porque “espiam” o barramento.
- A unidade de transferência e armazenamento de uma *cache* é denominada uma **linha de cache**.
- O protocolo de coerência de *cache* mais simples de todos é denominado **escrita direta**.

O protocolo MESI de coerência de *cache*

- O protocolo MESI é usado pelo Core i7 e por muitas outras CPUs para espionar o barramento.
- Cada entrada de *cache* pode estar em um dos quatro estados:
 1. Inválido – A entrada da *cache* não contém dados válidos.
 2. Compartilhado (*shared*) – múltiplas *caches* podem conter a linha; a memória está atualizada.
 3. Exclusivo – nenhuma outra *cache* contém a linha; a memória está atualizada.
 4. Modificado – a entrada é válida; a memória é inválida; não existem cópias.

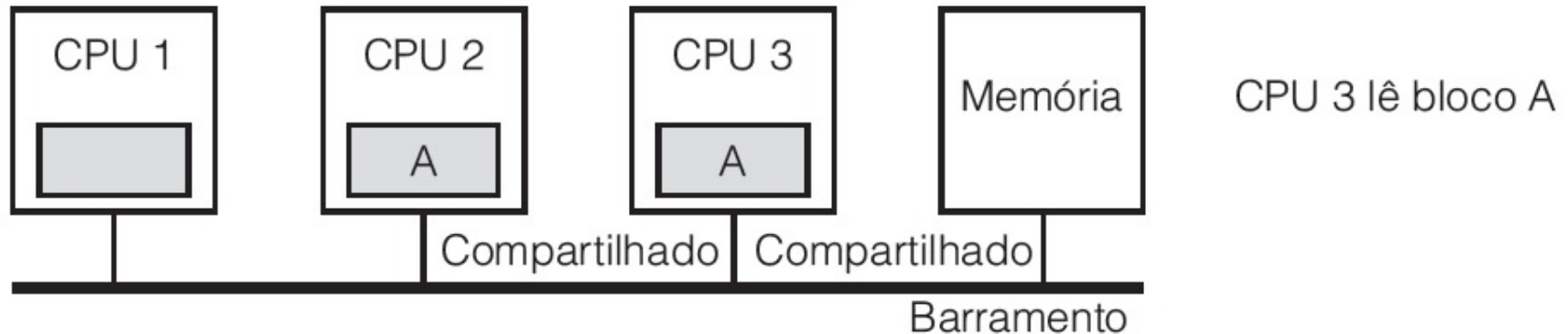
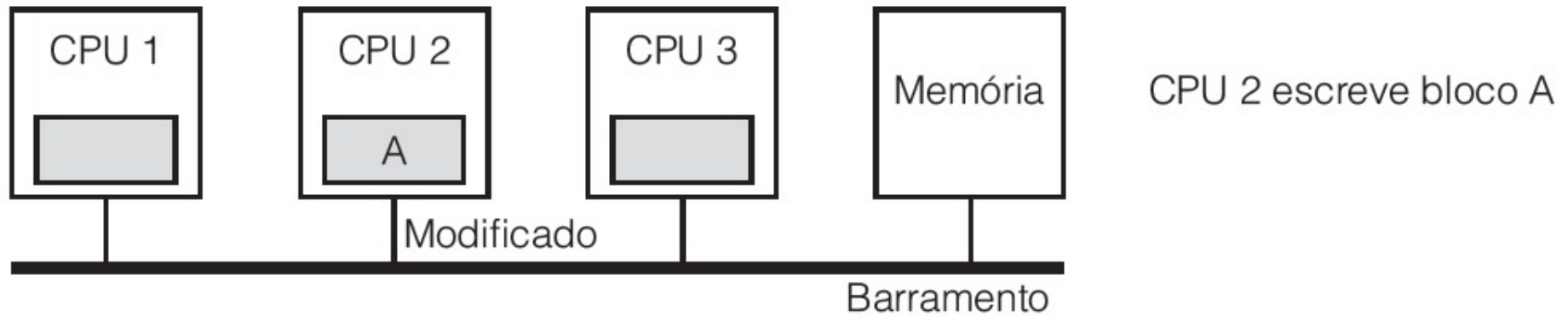
O protocolo MESI de coerência de *cache*

- Protocolo MESI de coerência de *cache*.



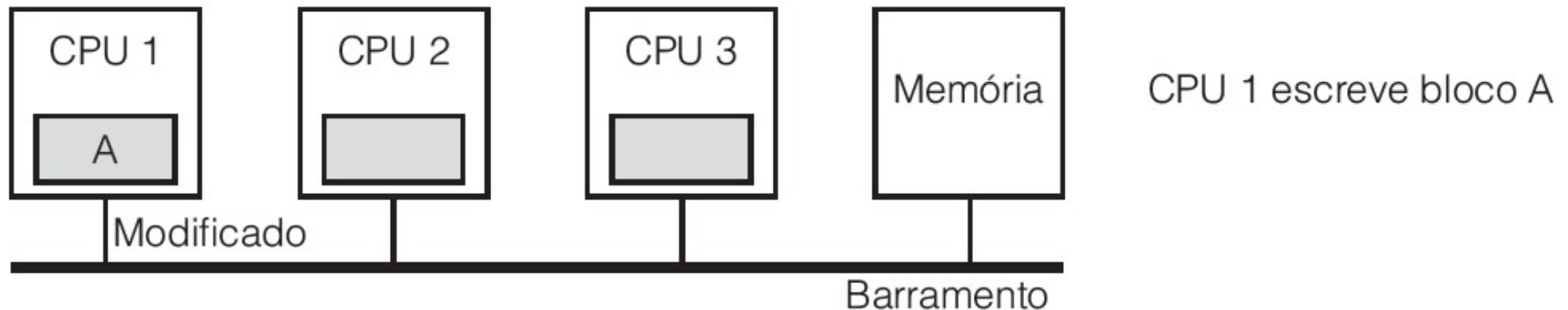
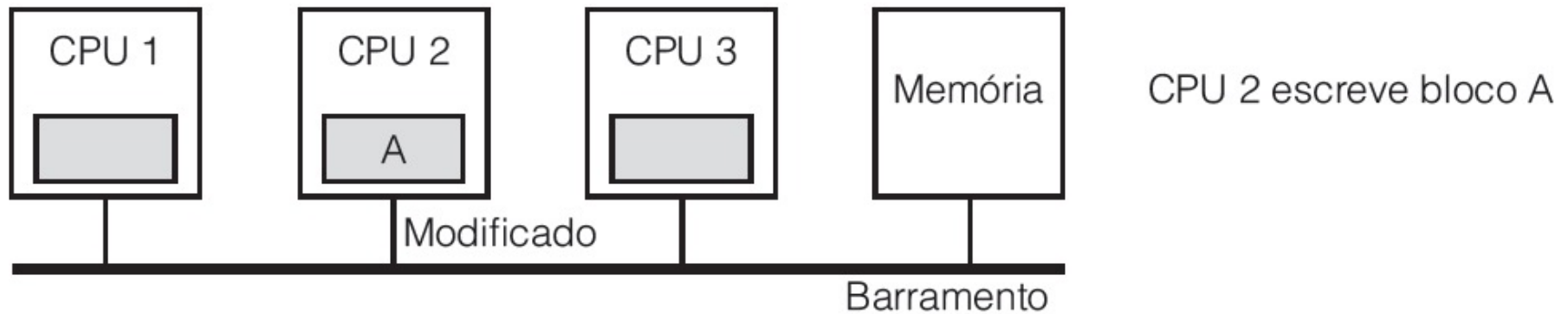
O protocolo MESI de coerência de *cache*

- Protocolo MESI de coerência de *cache*.



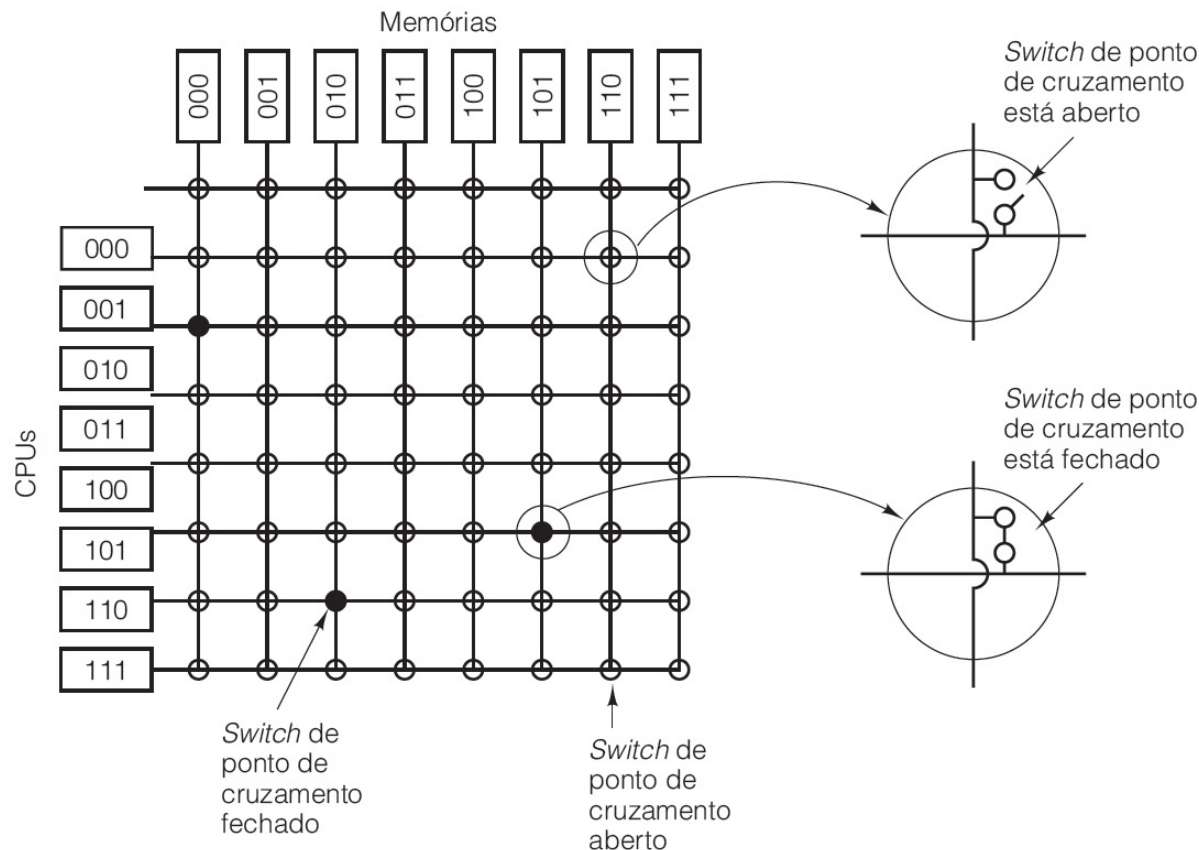
O protocolo MESI de coerência de *cache*

- Protocolo MESI de coerência de *cache*.



Multiprocessadores UMA que usam *switches crossbar*

- O circuito mais simples para conectar n CPUs a k memórias é o ***switch crossbar***.



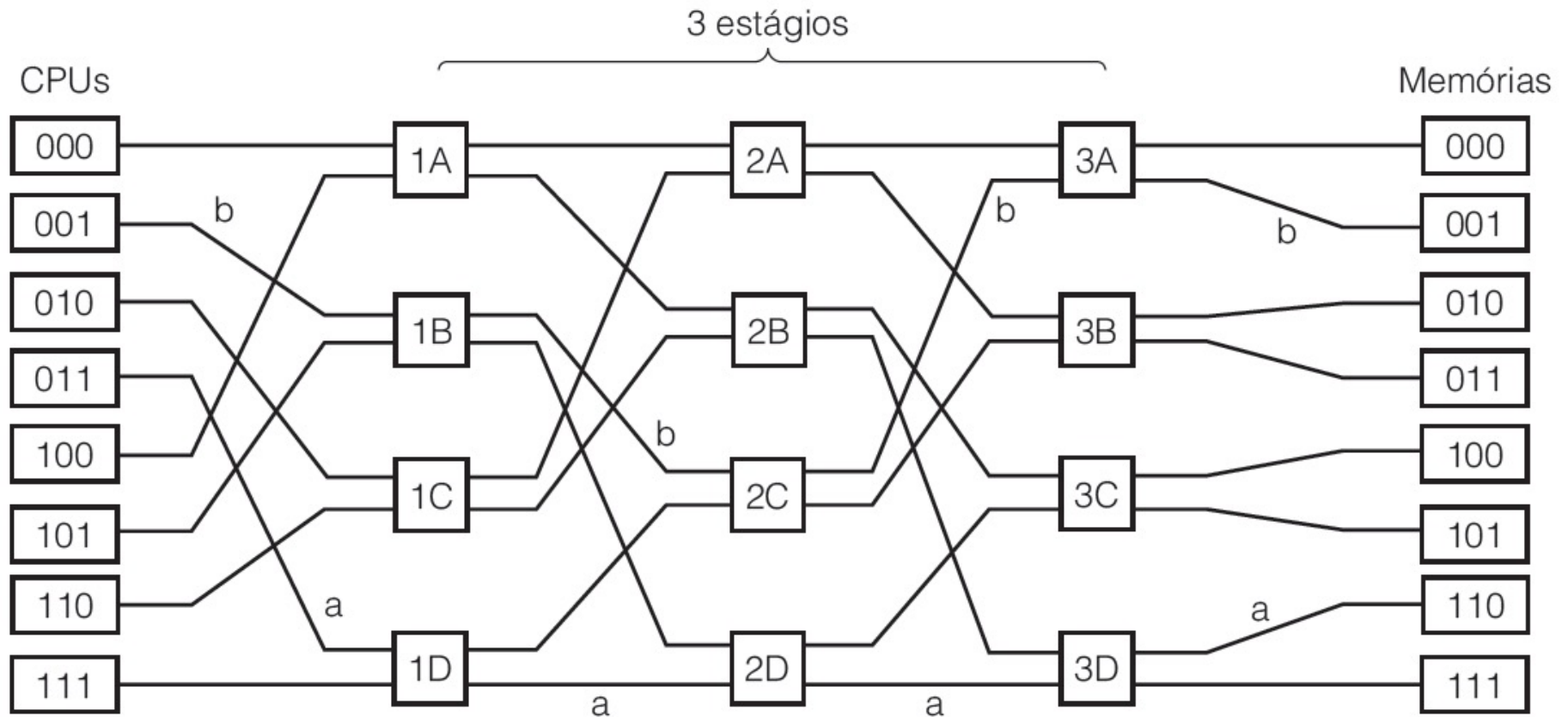
Multiprocessadores UMA que usam redes de comutação multiestágios

- Nossos *switches* 2×2 podem ser organizados de muitos modos para construir redes de comutação multiestágios maiores.



- Uma possibilidade é a rede ômega, classe econômica, sem supérfluos, ilustrada a seguir.
- O padrão de fiação da rede ômega costuma ser denominado **embaralhamento perfeito**.

Multiprocessadores UMA que usam redes de comutação multiestágios

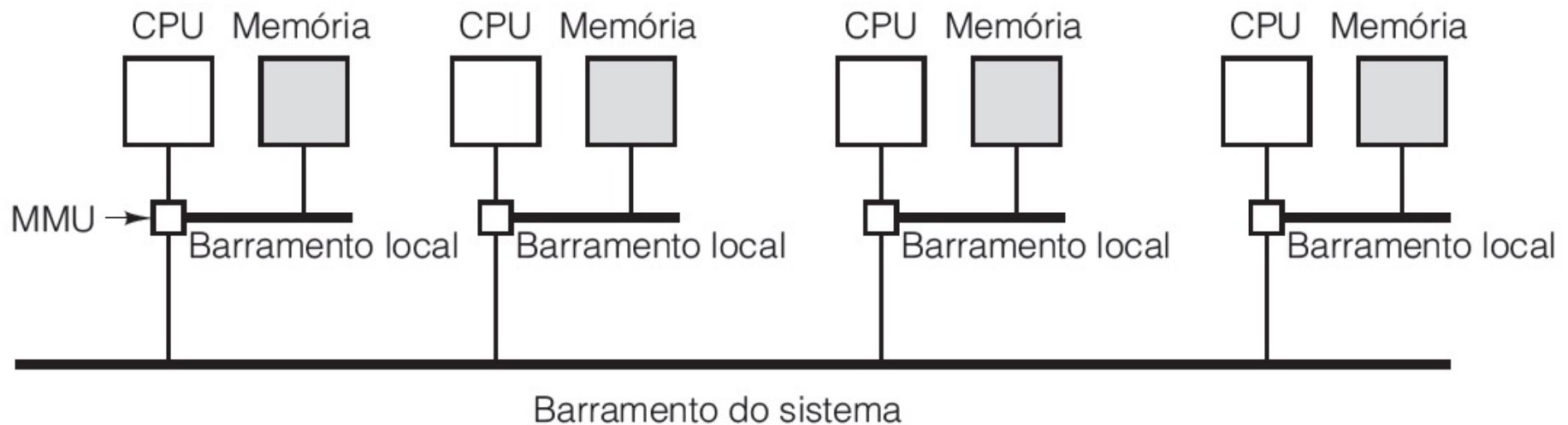


Multiprocessadores UMA que usam redes de comutação multiestágios

- Multiprocessadores **NUMA** fornecem um único espaço de endereço para todas as CPUs, porém, o acesso a módulos de memória locais é mais rápido do que o acesso a módulos remotos. As máquinas NUMA têm três características fundamentais:
 1. Há um único espaço de endereço visível a todas as CPUs.
 2. O acesso à memória remota é feito usando instruções LOAD e STORE.
 3. O acesso à memória remota é mais lento do que o acesso à memória local.

Multiprocessadores UMA que usam redes de comutação multistágios

- Máquina NUMA com dois níveis de barramentos. O Cm foi o primeiro multiprocessador a usar esse projeto.

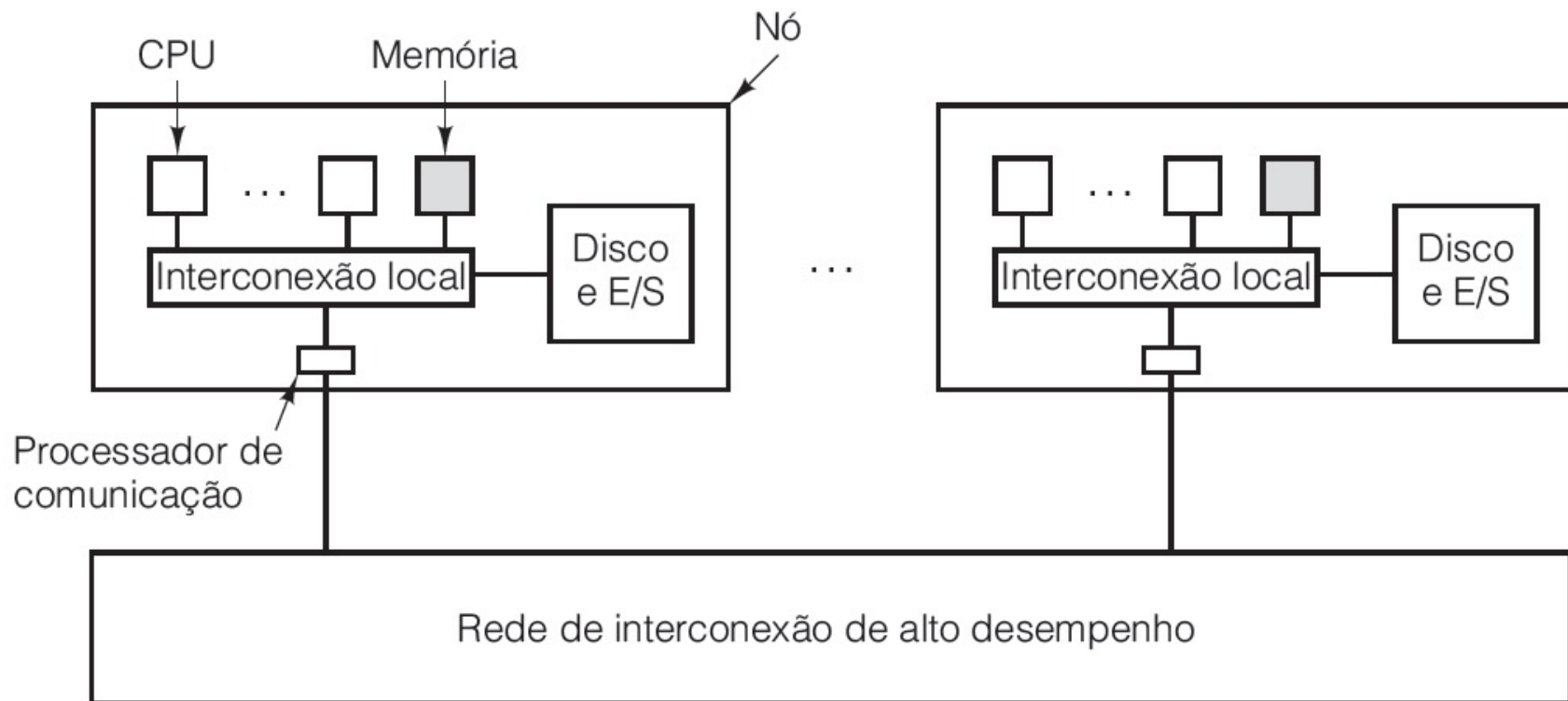


Multiprocessadores COMA

- **COMA** - as páginas não têm máquinas nativas fixas, como acontece em máquinas NUMA e CC-NUMA.
- Na verdade, as páginas não têm qualquer significado.
- Sistemas COMA introduzem dois novos problemas:
 1. Como as linhas de cache são localizadas?
 2. Quando uma linha é expurgada da memória, o que acontece se ela for a última cópia?

Multicomputadores de troca de mensagens

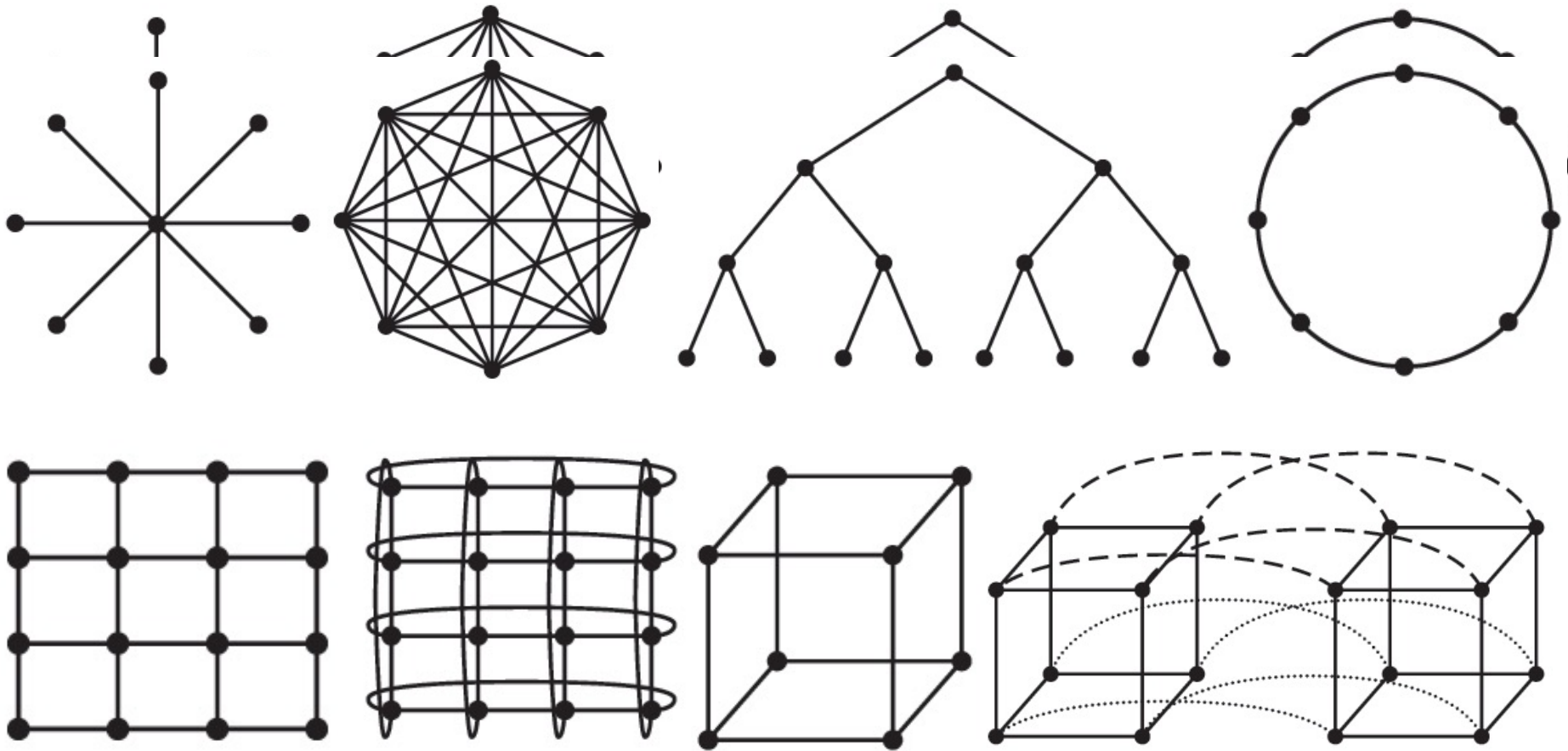
- Multicomputadores – computadores paralelos nos quais cada CPU tem sua própria memória privada, que não pode ser acessada diretamente por qualquer outra CPU.



Multicomputadores de troca de mensagens

- Na figura anterior, vemos que multicomputadores são mantidos juntos por redes de interconexão.
- A topologia de uma rede de interconexão descreve como os enlaces e os computadores são organizados, por exemplo, como um anel ou uma grade.
- Várias topologias são mostradas a seguir.
 - Estrela / Malha de interconexão completa
 - Árvore / Anel
 - Grade / Toro duplo
 - Cubo / Hipercubo 4D

Multicomputadores de troca de mensagens



MPPs – processadores maciçamente paralelos

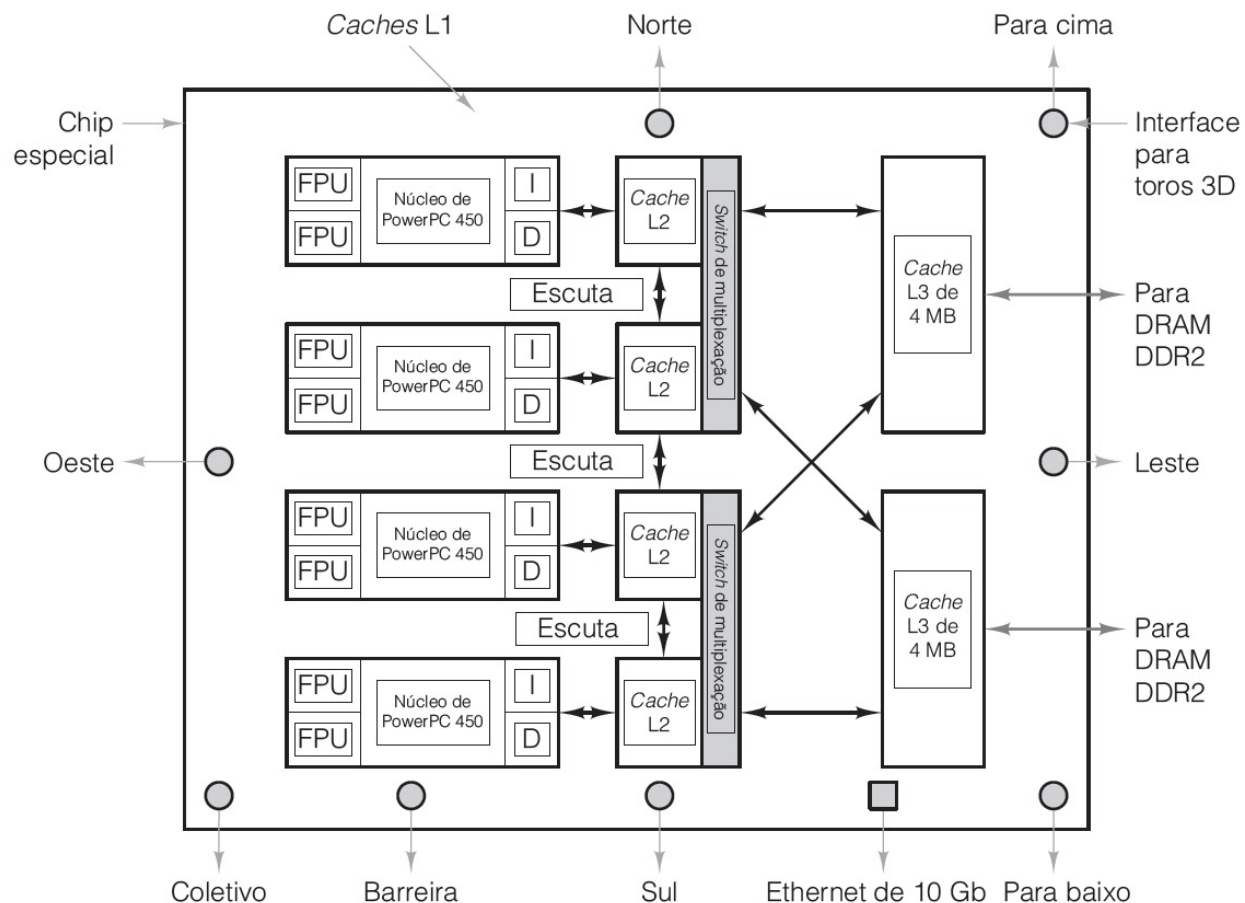
- MPPs são imensos supercomputadores de muitos milhões de dólares.
- Opções populares são o Pentium da Intel, a UltraSPARC da Sun e o PowerPC da IBM.
- O que caracteriza os MPPs é sua enorme capacidade de E/S e sua atenção com a tolerância à falha.
- Grandes MPPs sempre têm hardware e software especiais para monitorar o sistema, detectar falhas e recuperar-se delas facilmente.

MPPs – processadores maciçamente paralelos

BlueGene

O coração do
sistema

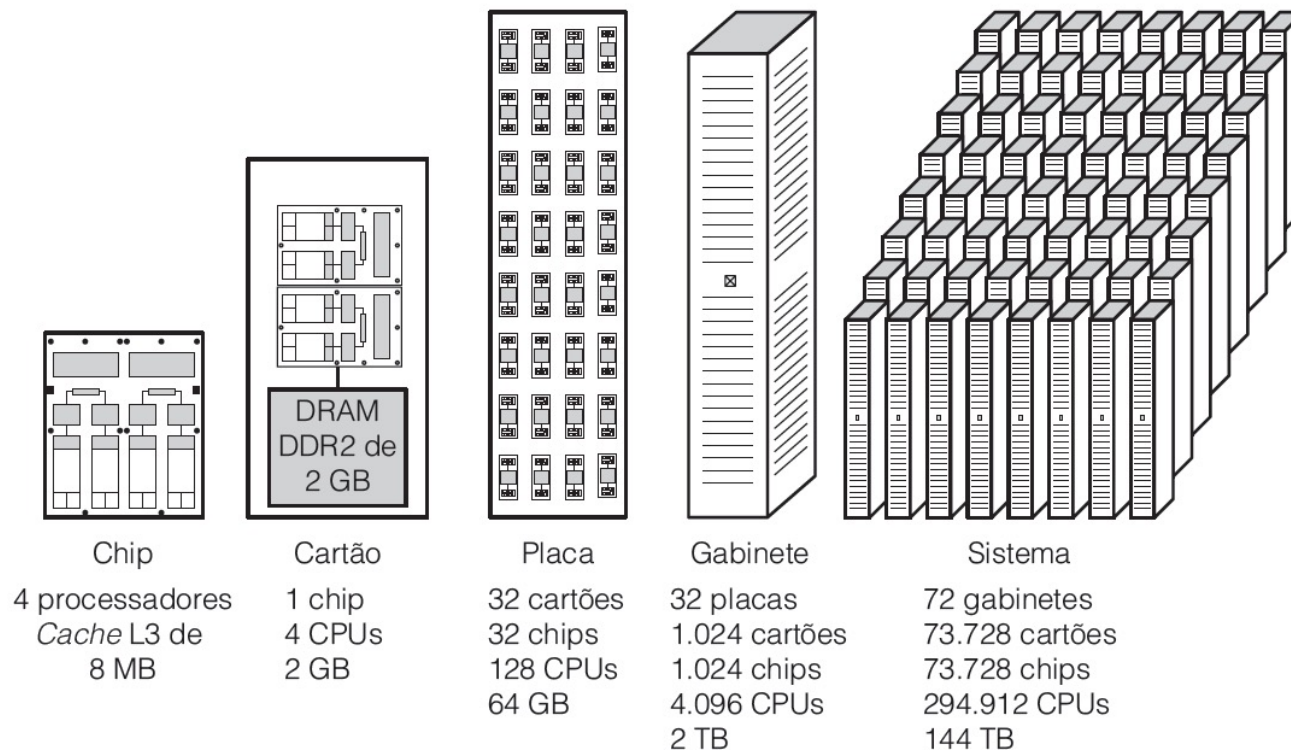
BlueGene/P é o
chip especial de
nó.



MPPs – processadores maciçamente paralelos

BlueGene

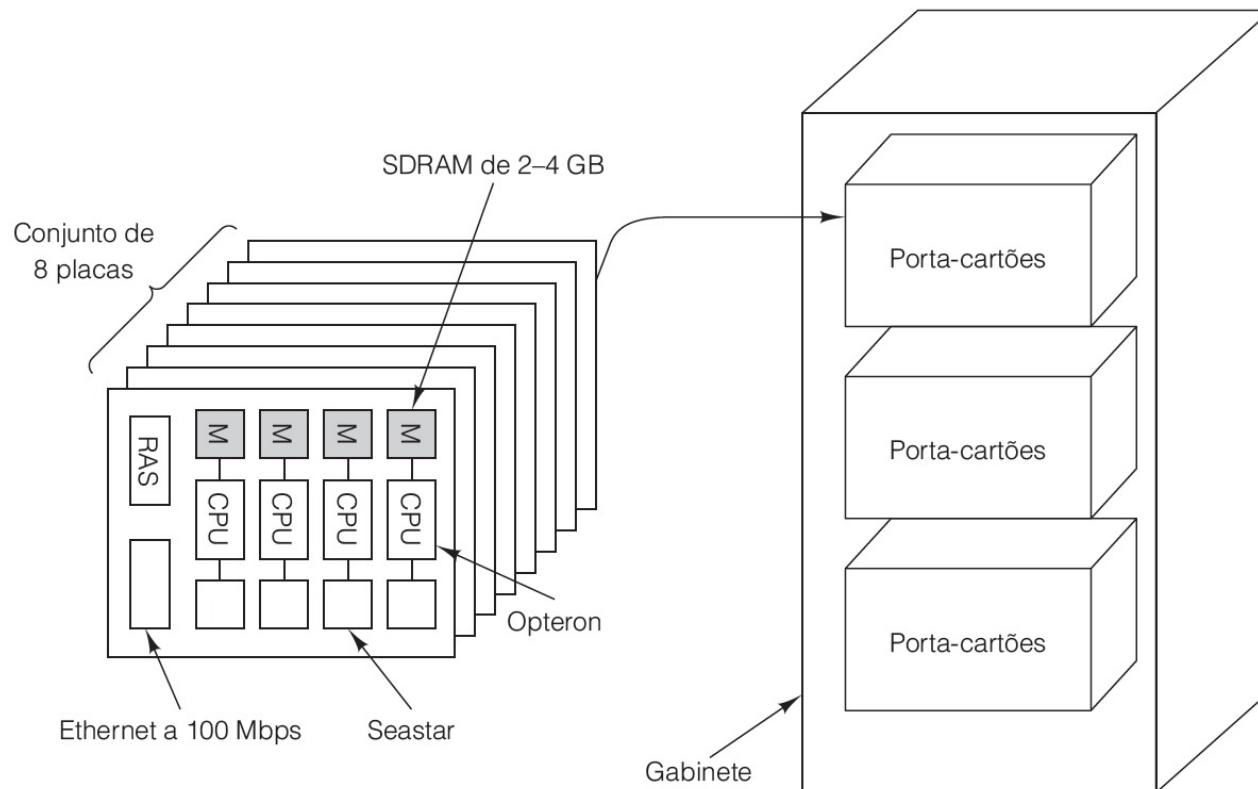
- BlueGene/P. Chip. Cartão. Placa. Gabinete. Sistema.



MPPs – processadores maciçamente paralelos

Red Storm

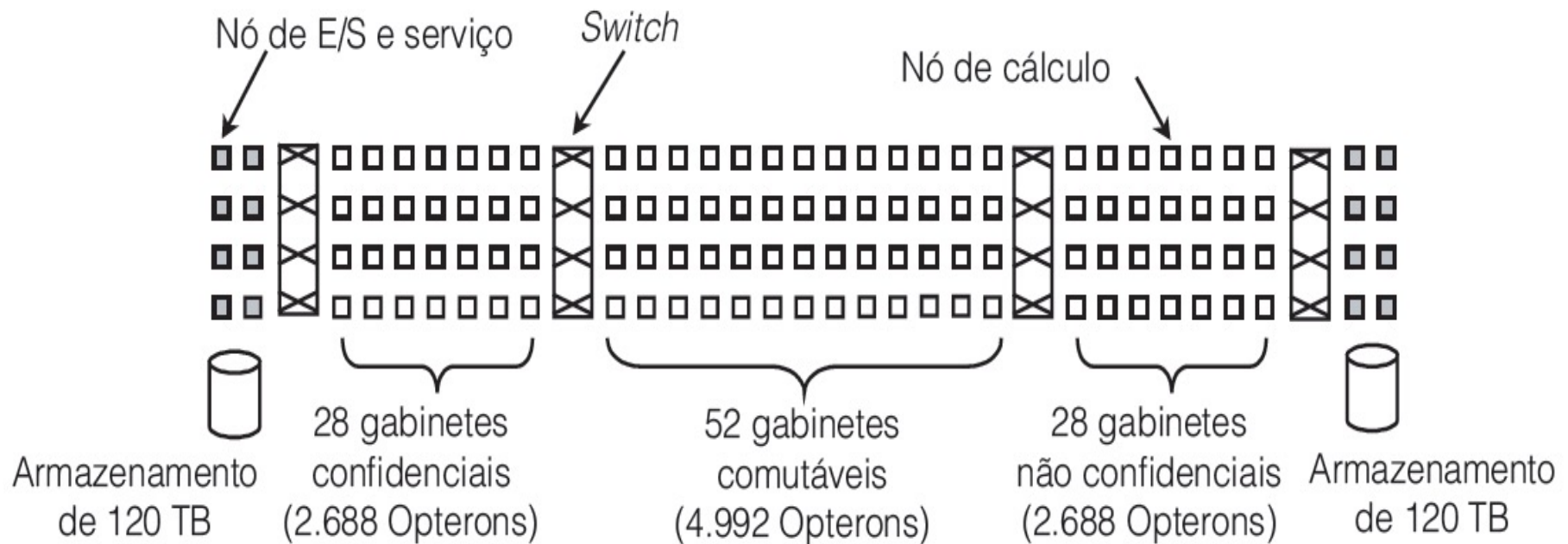
- Pacote de componentes do Red Storm.



MPPs – processadores maciçamente paralelos

Red Storm

- Sistema Red Storm visto de cima.

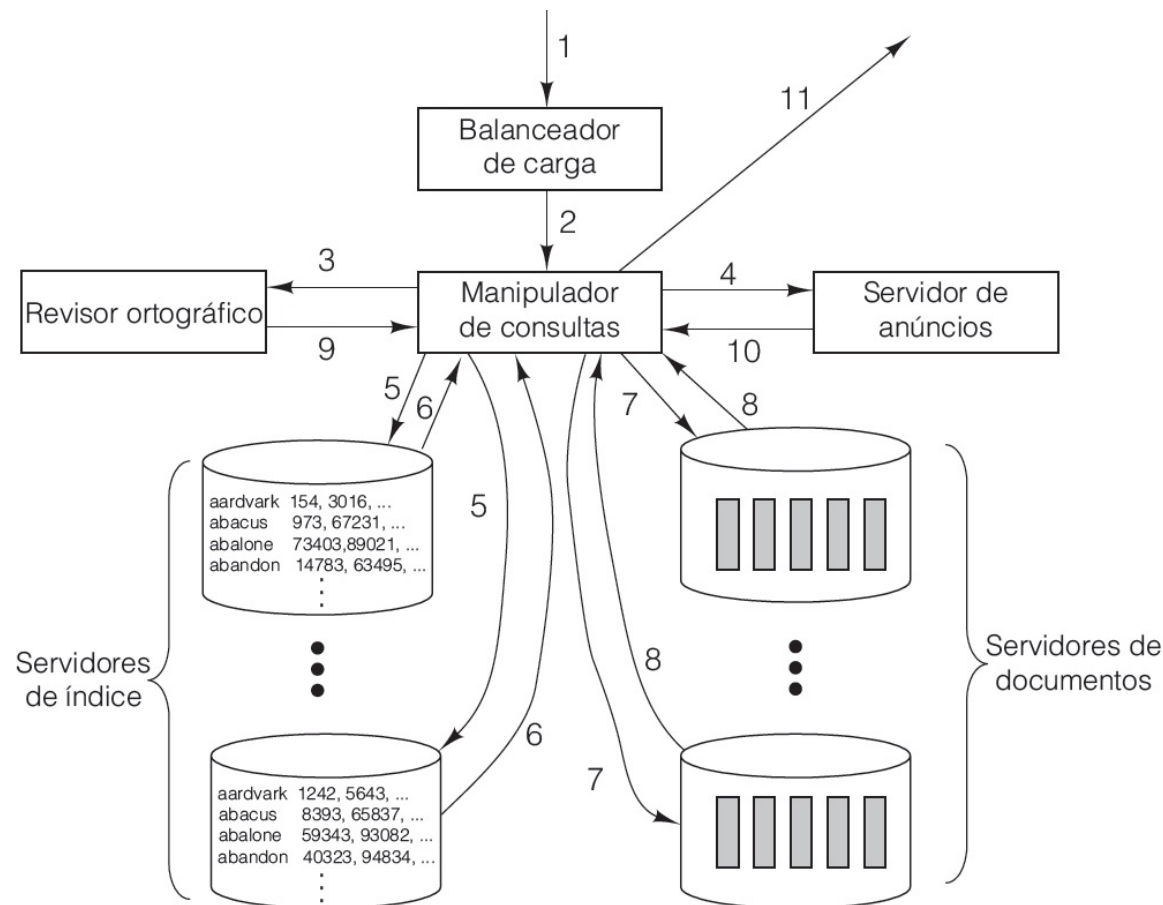


Computação de cluster

- Consiste em centenas de milhares de PCs ou estações de trabalho conectadas por uma placa de rede disponível no mercado.
- Embora existam muitos tipos de clusters, há duas espécies que dominam: o centralizado e o descentralizado.
- De modo geral, os clusters são conjuntos pequenos, na faixa de uma dúzia a talvez 500 PCs.
- Contudo, também é possível construir clusters muito grandes com PCs de prateleira.

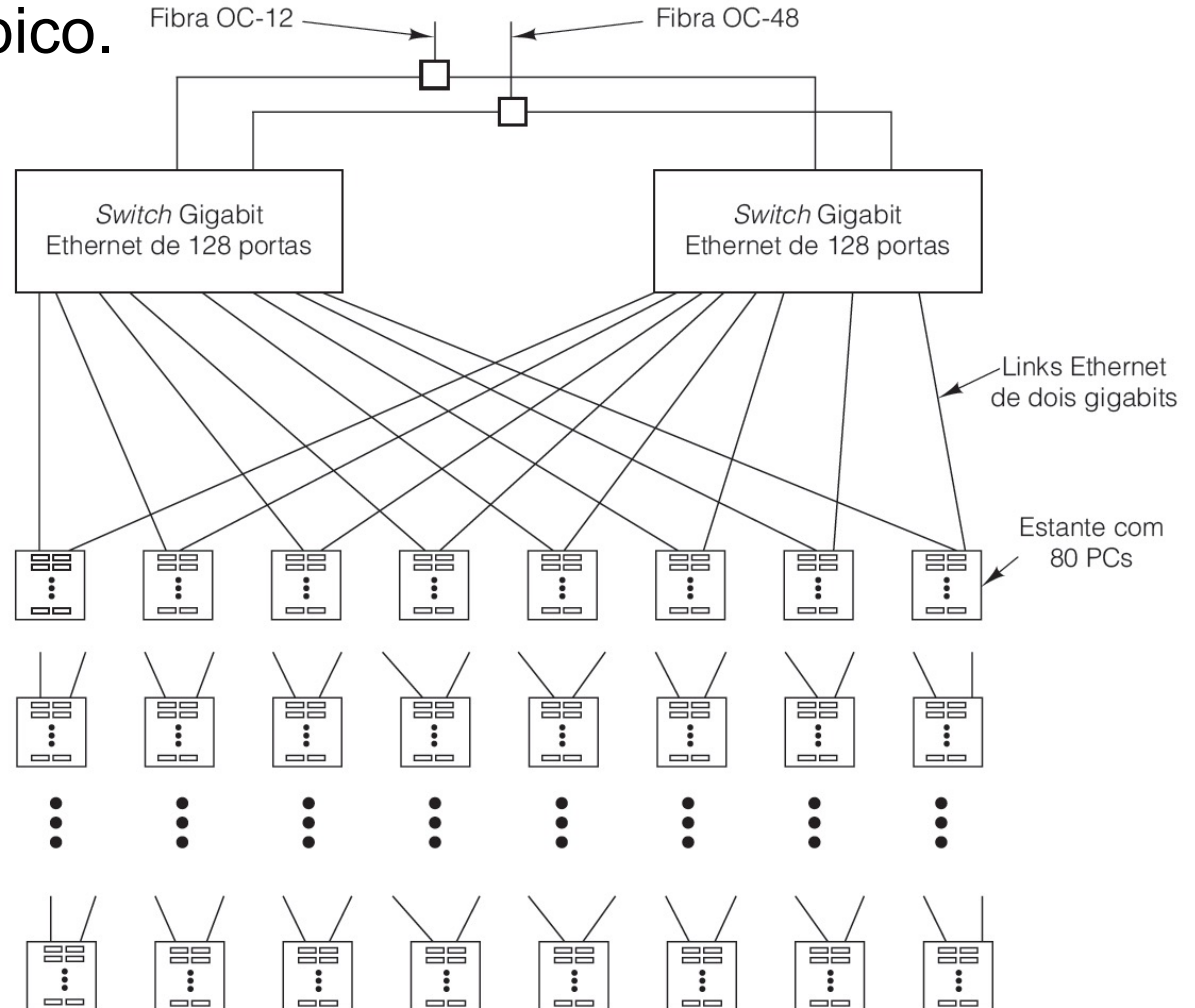
Google

- Processamento de uma consulta no Google.



Google

- *Cluster Google típico.*

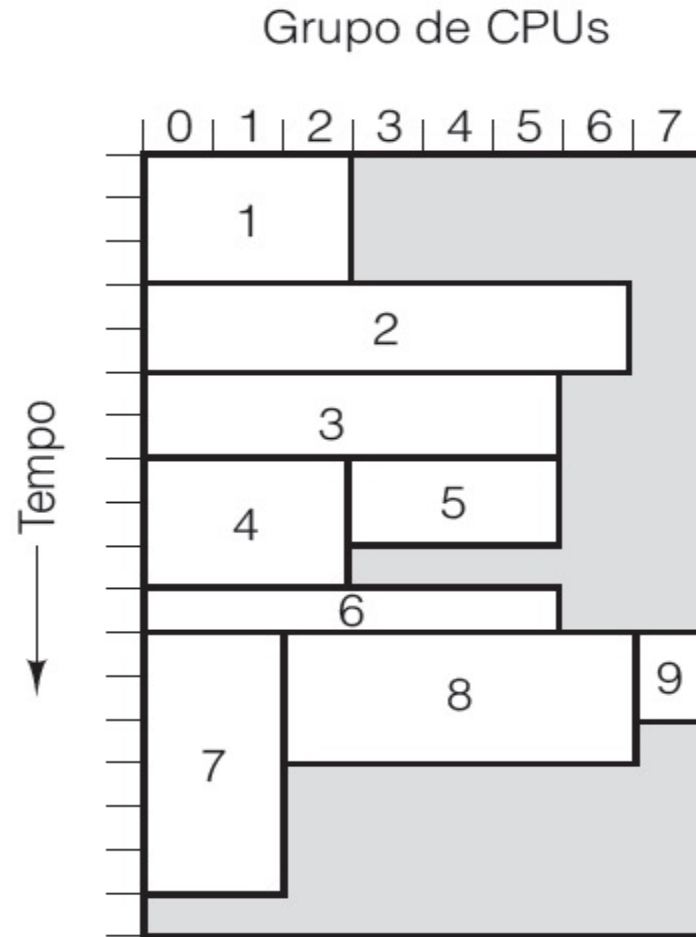


Software de comunicação para multicomputadores

- A maioria dos sistemas de troca de mensagens fornece duas primitivas (em geral, chamadas de biblioteca), *send* e *receive*, mas diversos tipos de semânticas são possíveis.
- As três variantes principais são:
 1. Troca síncrona de mensagens.
 2. Troca de mensagens por *buffers*.
 3. Troca de mensagens sem bloqueio.

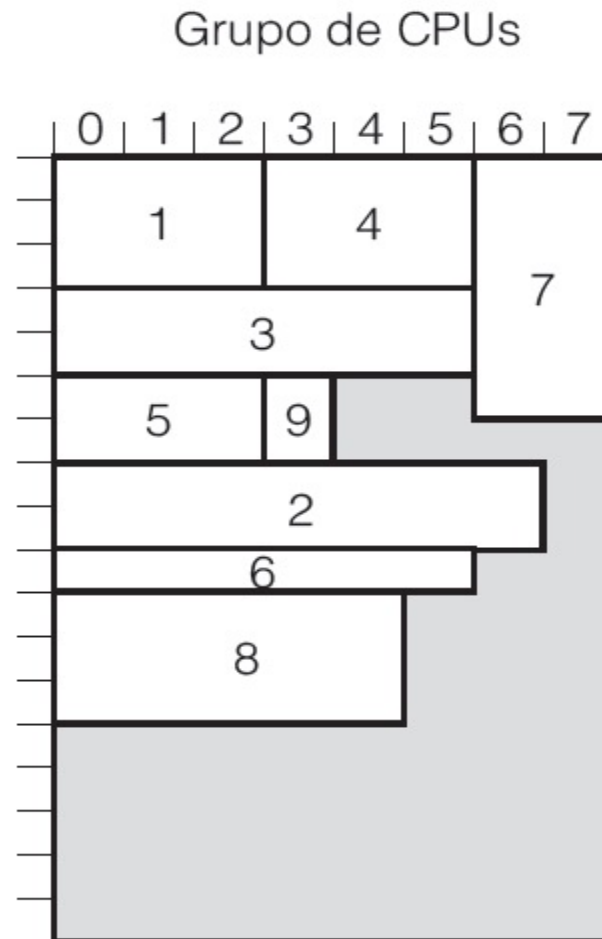
Escalonamento

- Escalonamento de um cluster. FIFO.



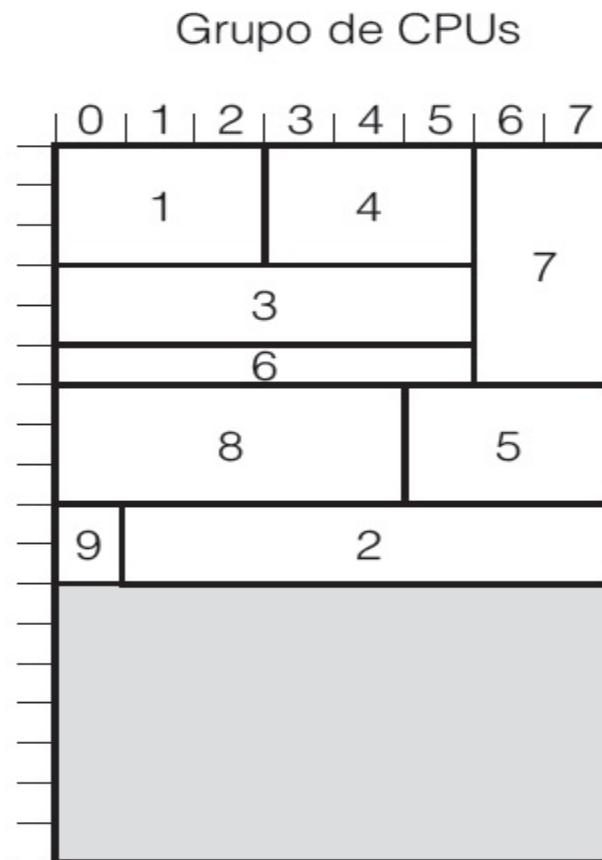
Escalonamento

- Escalonamento de um cluster. Sem bloqueio de cabeça de fila.



Escalonamento

- Escalonamento de um cluster. Lajotas. As áreas sombreadas indicam CPUs ociosas.



Memória compartilhada distribuída

- **DSM** - um conjunto de CPUs em um multicomputador compartilha um espaço de endereço virtual paginado.
- Na versão mais simples, cada página é mantida na RAM de exatamente uma CPU.
- Visto que só existe uma cópia da página, ela ficará em constante ir e vir, uma situação conhecida como **falso compartilhamento**.
- Uma otimização feita em Treadmarks é mapear inicialmente cada página que pode ser escrita, em modo somente de leitura.

Linda

- Fornece processos em várias máquinas com uma memória compartilhada distribuída com alto grau de estruturação.
- O conceito unificador fundamental de Linda é o de um **espaço de tuplas** abstrato, que é global para o sistema inteiro e acessível a todos os seus processos.
- Três tuplas Linda.

("abc", 2, 5)

("matrix-1", 1, 6, 3.14)

("family", "is sister", Carolyn, Elinor)

Desempenho

- As velocidades de CPU e E/S são as mesmas de um uniprocessador.
- A latência de ida e volta é o tempo que leva para a CPU enviar um pacote e obter uma resposta.
- A latência que interessa é a de pacotes mínimos, muitas vezes uma única palavra ou uma pequena linha de *cache*.
- O número de bytes/s que o sistema pode mover é crítico para o desempenho.
- Há várias métricas para largura de banda.

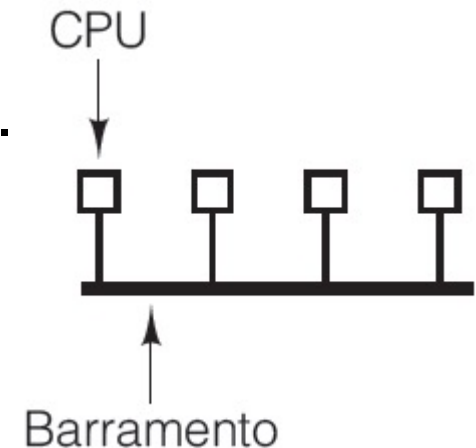
Desempenho

- Para os usuários, a métrica fundamental é mais velocidade:
 - quanto mais depressa um programa executa em um sistema de n processadores em comparação com um sistema com um só processador.
- O aumento de velocidade é apenas o tempo de execução do programa original, T , dividido pelo novo tempo de execução:

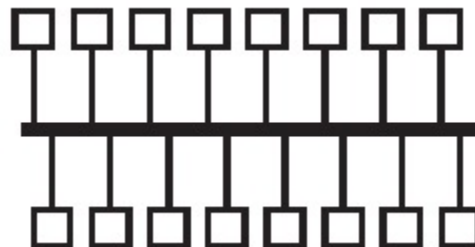
$$\text{Aumento de velocidade} = \frac{n}{1 + (n - 1) f}$$

Desempenho

- O modo mais direto de melhorar o desempenho é adicionar CPUs ao sistema.
- Sistema de 4 CPUs com um barramento.

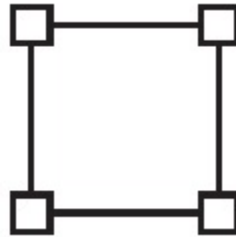


- Sistema de 16 CPUs com um barramento.

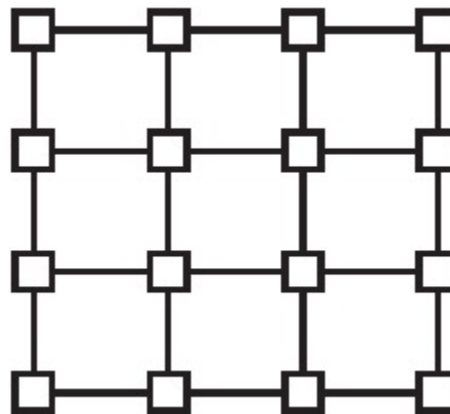


Desempenho

- Sistema de 4 CPUs em grade.



- Sistema de 16 CPUs em grade.



Computação em grade

- O objetivo da grade é proporcionar infraestrutura técnica para permitir que um grupo de organizações que compartilham uma mesma meta forme uma **organização virtual**.
- A grade é inerentemente multilateral, com muitos participantes de mesmo *status*.
- A grade precisa ter acesso a uma ampla variedade de recursos.
- Um modo de modelar a grade é a hierarquia em camadas.

Computação em grade

- Camadas da grade.

Camada	Função
Aplicação	Aplicações que compartilham recursos gerenciados de modos controlados
Coletiva	Descoberta, corretagem, monitoração e controle de grupos de recursos
De recursos	Acesso seguro e gerenciado a recursos individuais
Base	Recursos físicos: computadores, armazenamento, redes, sensores, programas e dados

Computação em grade

- A comunidade das grades criou uma organização, a Global Grid Forum, para gerenciar o processo de padronização.
- Ela criou uma estrutura denominada OGSA.
- Os serviços que estão atualmente em fase de padronização pertencem a oito categorias gerais:
 1. Serviços de infraestrutura (habilitar comunicação entre recursos).
 2. Serviços de gerenciamento de recursos (reserva e distribuição de recursos).

Computação em grade

3. Serviços de dados (mover e replicar dados para onde forem necessários).
4. Serviços de contexto (descrever recursos necessários e políticas de utilização).
5. Serviços de informação (obter informações sobre disponibilidade de recursos).
6. Serviços de autogerenciamento (suportar uma qualidade de serviço declarada).
7. Serviços de segurança (impor políticas de segurança).
8. Serviços de gerenciamento de execução (gerenciar fluxo de trabalho).