

Adaptação dos slides dos livros:

Capítulo 3

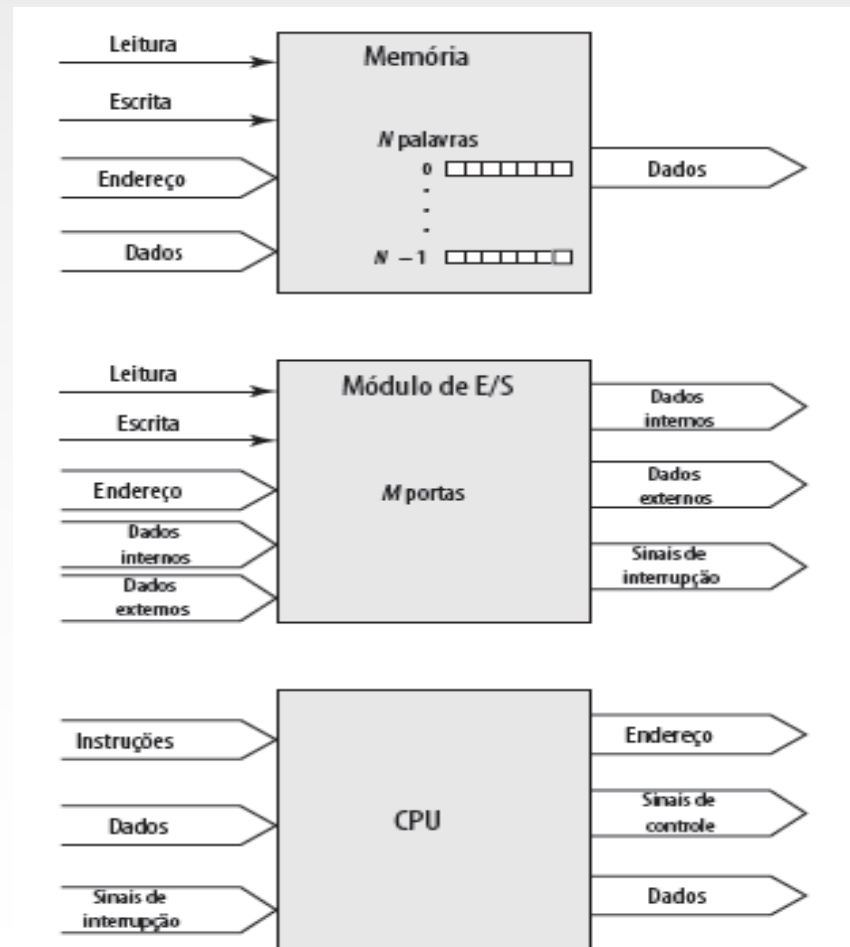
Visão de alto nível da função e interconexão do computador



Conectando

- Todas as unidades devem ser conectadas.
- Tipo de conexão diferente para tipo de unidade diferente.
 - Memória.
 - Entrada/saída.
 - CPU.

Módulos do computador



Conexão de memória

- Recebe e envia dados.
- Recebe endereços (de locais).
- Recebe sinais de controle:
 - Leitura.
 - Escrita.
 - Temporização.

Conexão de entrada/saída

- Semelhante à memória do ponto de vista do computador.
- Saída:
 - Recebe dados do computador.
 - Envia dados a periféricos.
- Entrada:
 - Recebe dados de periféricos.
 - Envia dados ao computador.

Conexão de entrada/saída

- Recebe sinais de controle do computador.
- Envia sinais de controle aos periféricos.
 - P.e., girar disco.
- Recebe endereços do computador.
 - P.e., número de porta para identificar periférico.
- Envia sinais de interrupção (controle).

Conexão da CPU

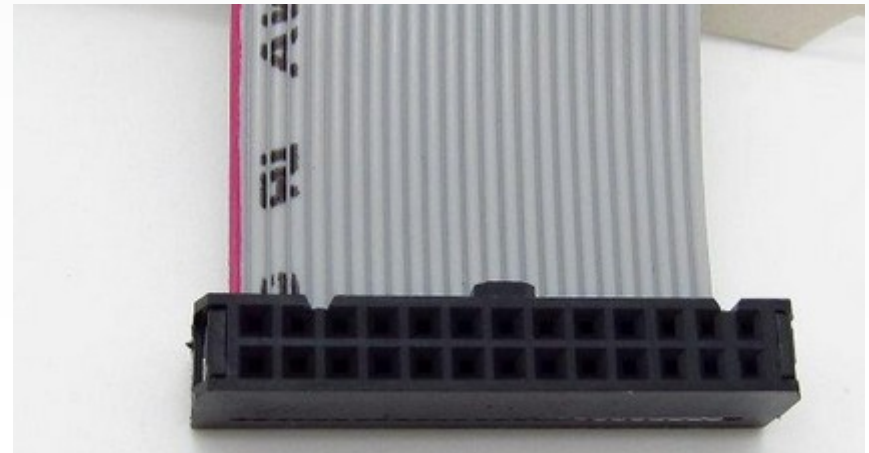
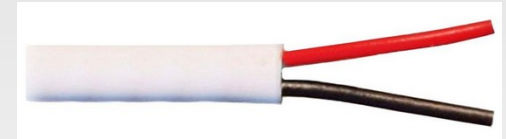
- Lê instruções e dados.
- Escreve dados (após processamento).
- Envia sinais de controle a outras unidades.
- Recebe (e atua sobre) interrupções.

Barramentos

- Existem diversos sistemas de interconexão possíveis.
- Estrutura de barramento único e múltiplo são mais comuns.
- P.e., barramento de Controle/Endereço/Dados (PC).
- P.e., Unibus (DEC-PDP).

Entendendo a Comunicação Paralela

- Comunicação serial:
 - 1 bit enviado por pulso de clock
 - Eficaz em distâncias longas
 - Exemplo: Código Morse
- Comunicação Paralela
 - N bits enviados por pulso de clock
 - Eficaz somente em curtas distâncias



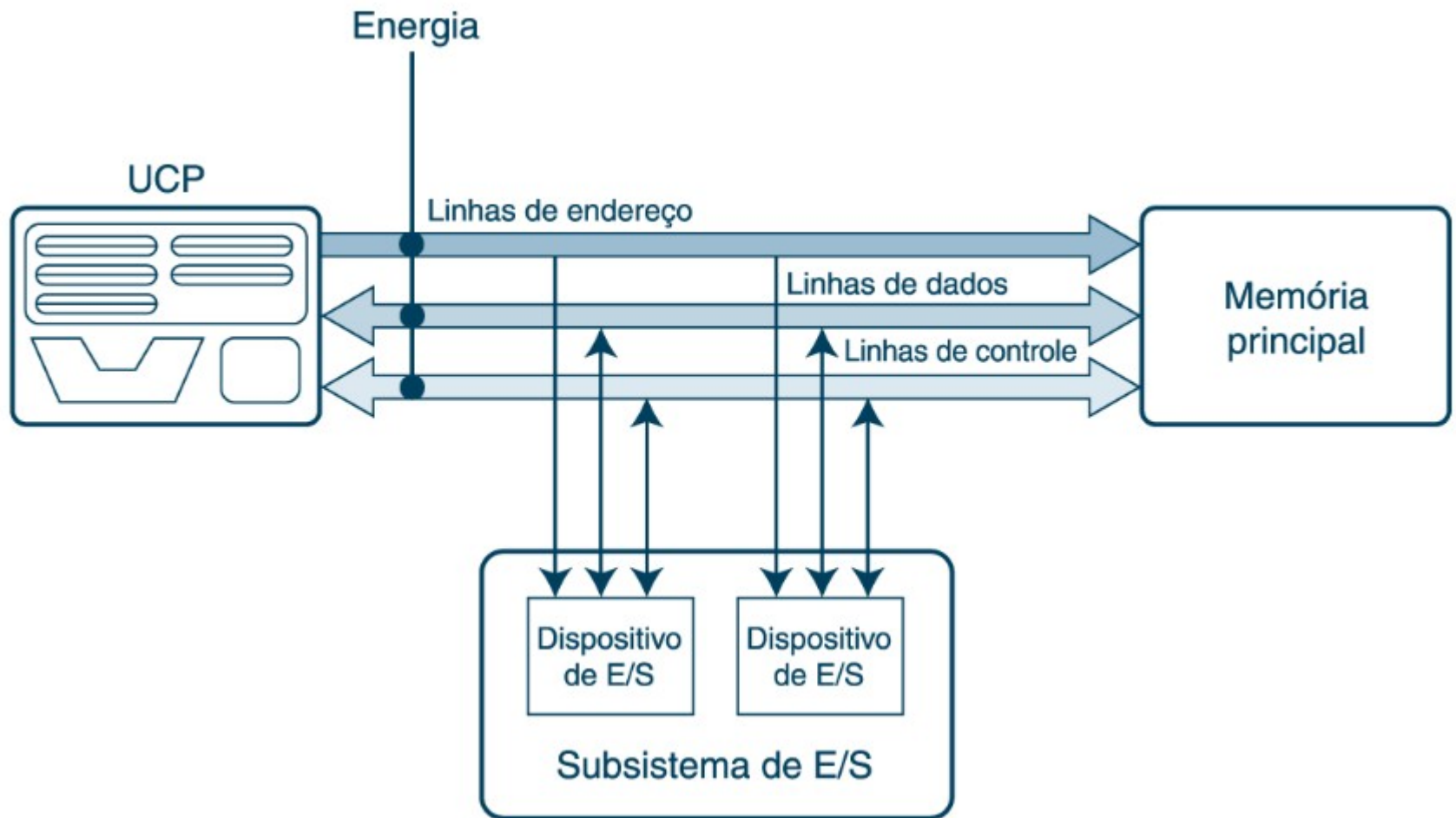
O que é um barramento?

Definição: Um barramento é um conjunto de fios que atuam como um caminho de dados comum, porém compartilhado, para conectar vários subsistemas dentro do sistema.

Normalmente, broadcast.

- Frequentemente agrupado
- Uma série de canais em um barramento.
- P.e., barramento de dados de 32 bits são 32 canais de bits separados.
- Linhas de potência podem não ser mostradas.

Barramento Típico



Linhas de Barramento

- Em um computador típico temos 3 linhas de barramentos distintos:
 - Barramento de Dados
 - Barramento de Endereços
 - Barramento de Controle

Barramento de dados

- Dedicados à transmissão de dados.
- Lembre-se de que não existe diferença entre “dados” e “instruções” neste nível.
- Largura é um determinante fundamental do desempenho.
8, 16, 32, 64 bits.

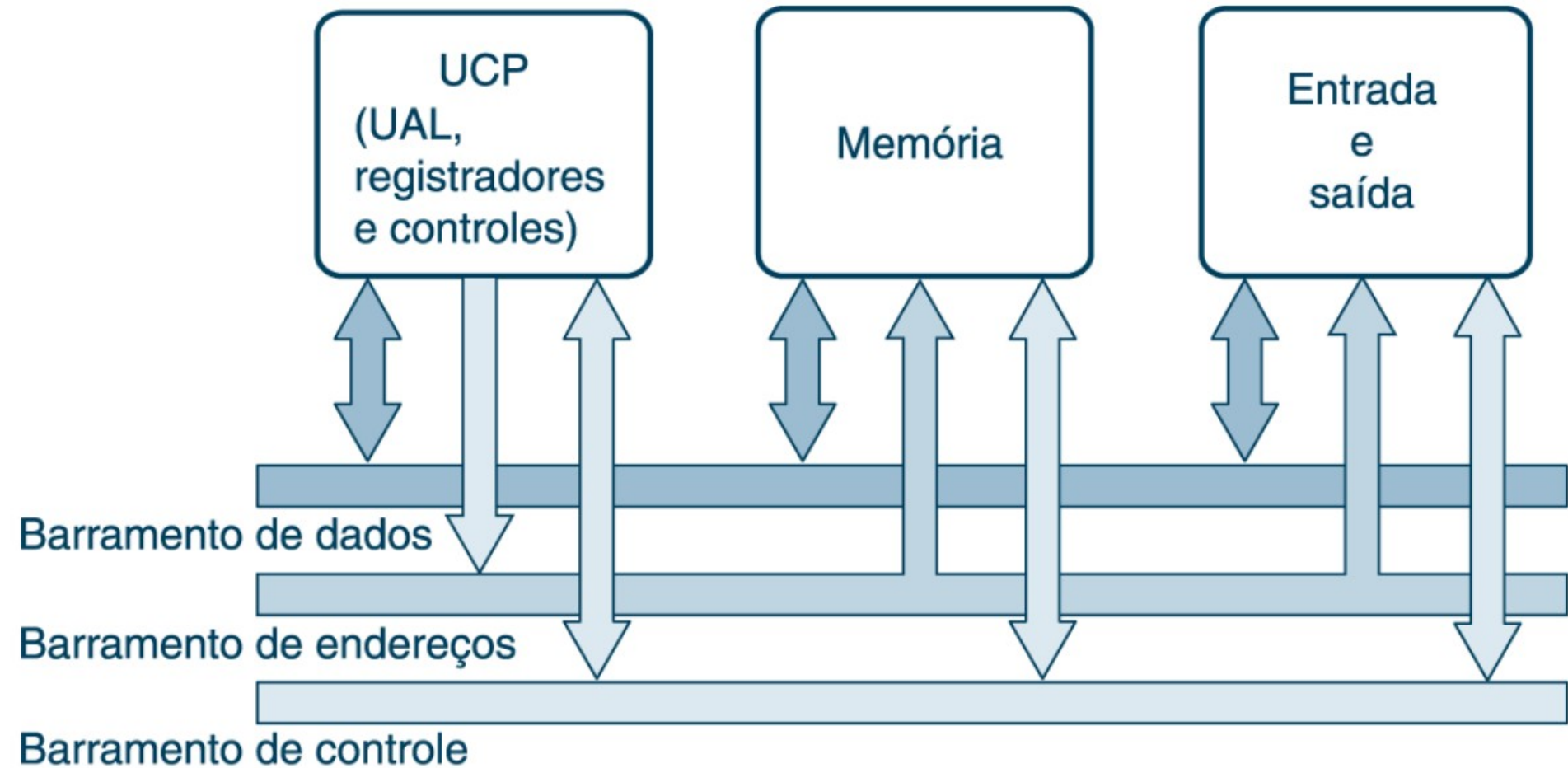
Barramento de endereço

- Identifica a posição em que o dado deve ser lido ou escrito.
- P.e., CPU precisa ler uma instrução (dados) de determinado local na memória.
- Largura do barramento determina capacidade máxima da memória do sistema.
- P.e., 8080 tem barramento de endereço de 16 bits gerando espaço de endereços de 64k.

Barramento de controle

- Indica qual dispositivo tem permissão para utilizar o barramento e qual a finalidade (leitura ou escrita em memória ou dispositivo de I/O)
- Informação de controle e temporização:
 - Sinal de leitura/escrita de memória.
 - Solicitação de interrupção.
 - Sinais de clock.

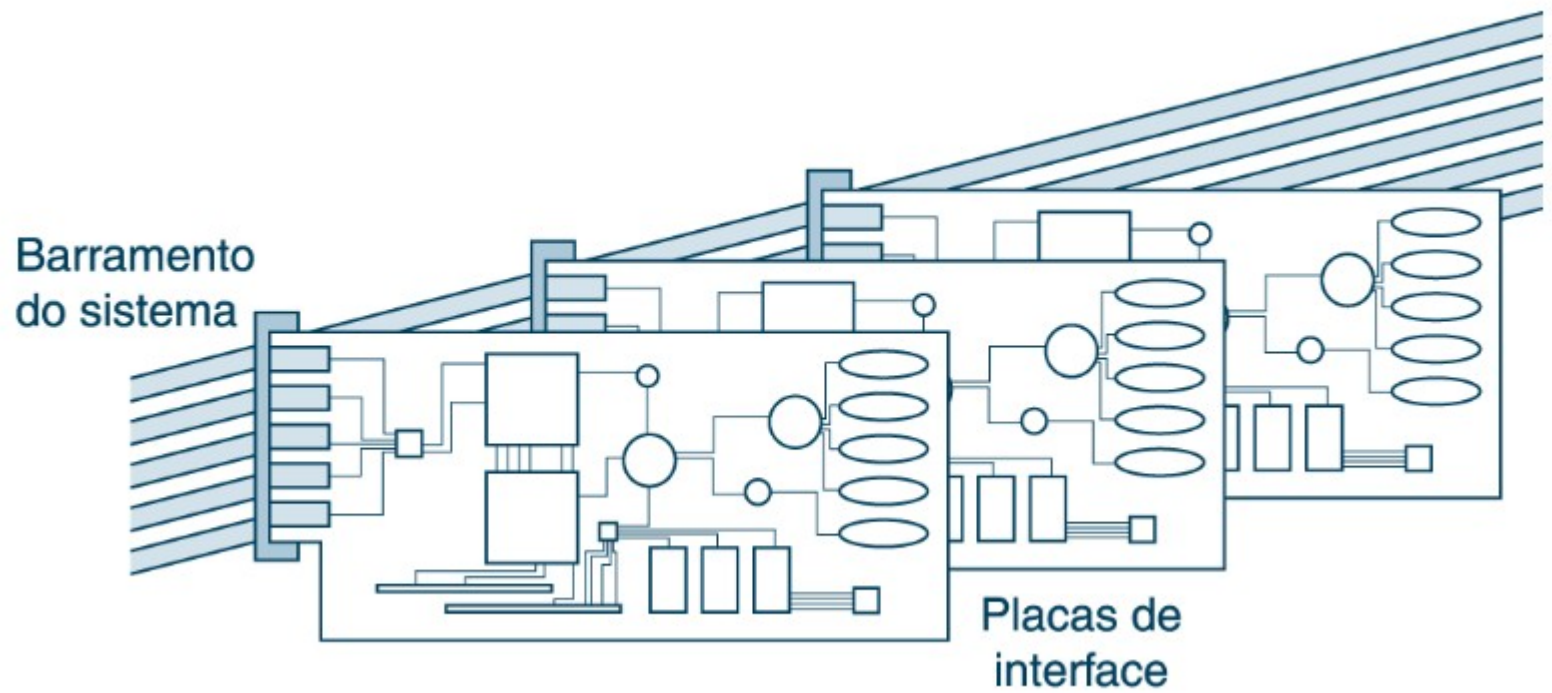
Esquema de interconexão de barramento



Um ônibus grande e amarelo?

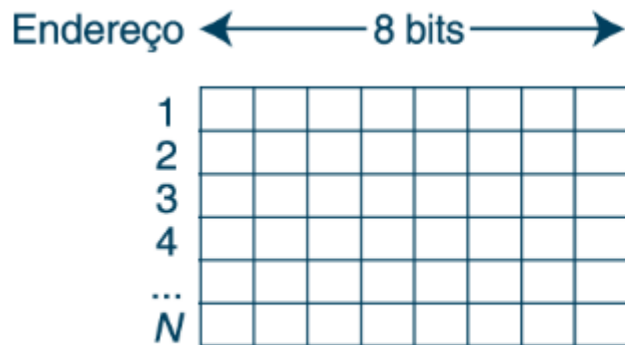
- Como os barramentos se parecem?
 - Linhas paralelas em placas de circuito.
 - Cabos de fita.
 - Conectores em tira nas placas mãe.
 - P.e., PCI.
 - Conjuntos de fios.

Barramento Traseiro I/O

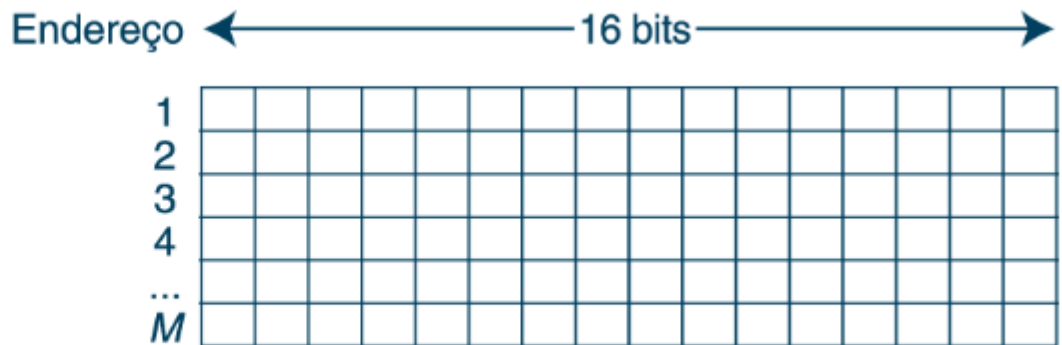


Problemas do barramento único

- Considerando que só existe processador e memória principal em um computador.
- A frequência do barramento é de 80 MHz.
- Qual a taxa de transferência no caso (a)? E no caso (b)?



(a)

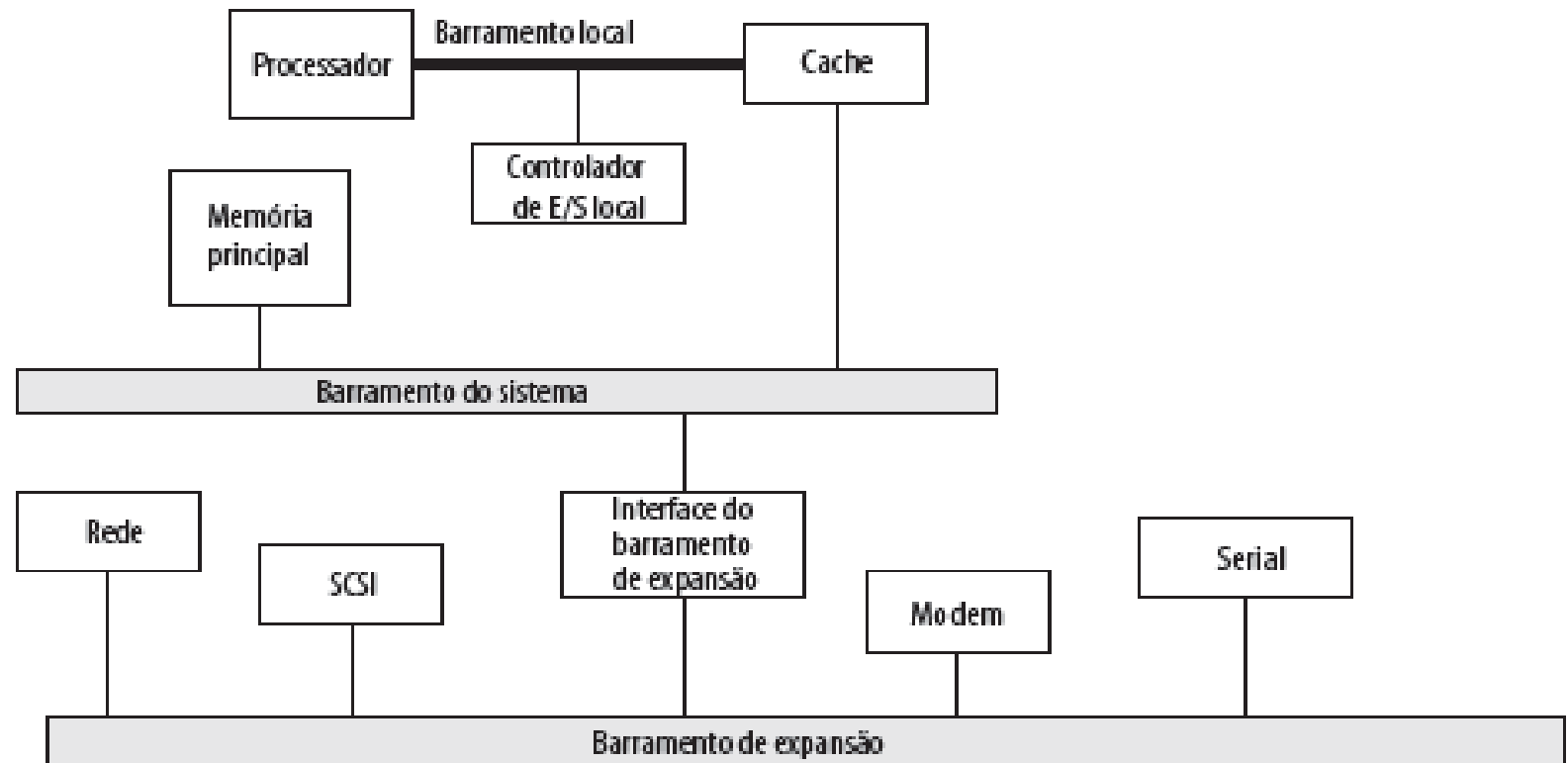


(b)

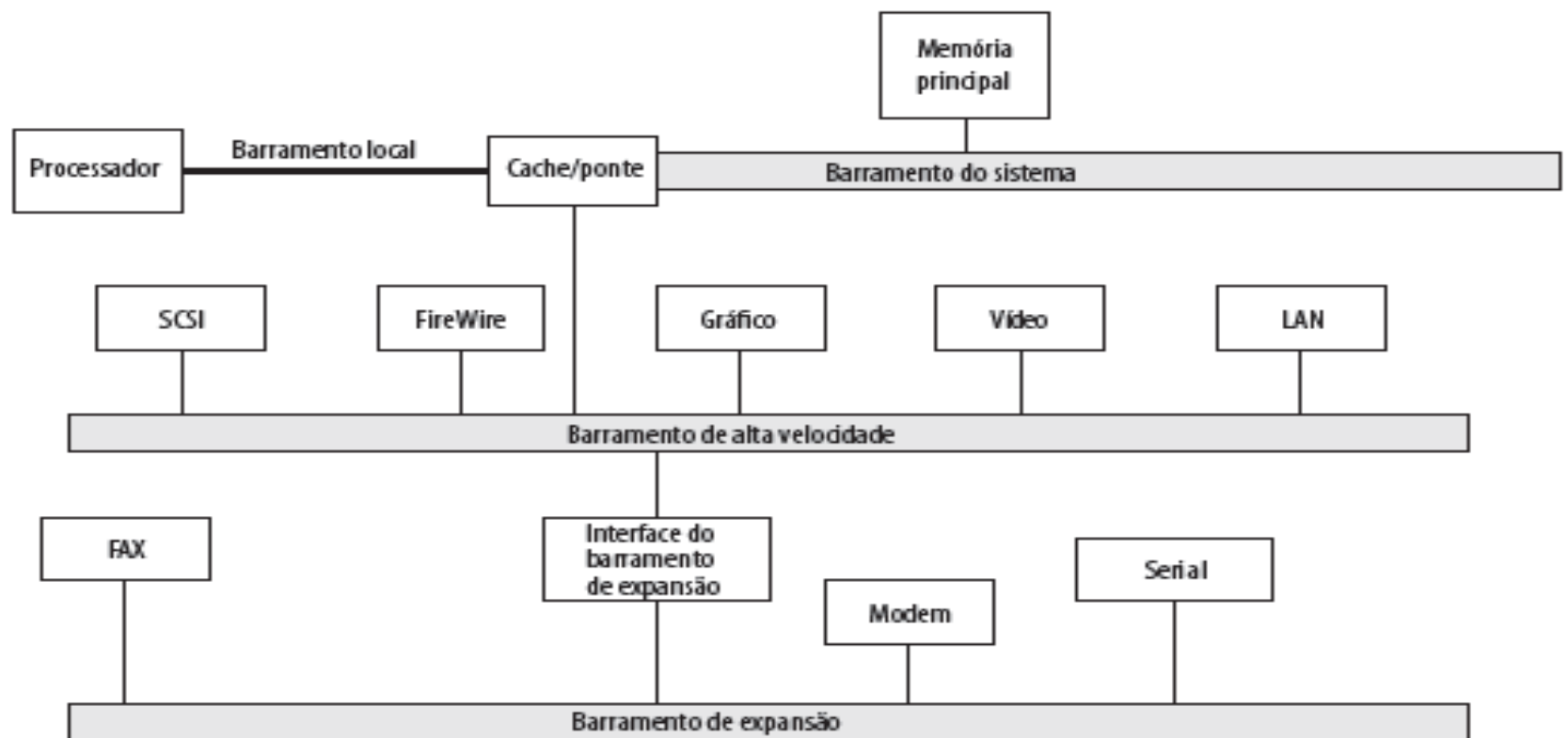
Problemas do barramento único

- Muitos dispositivos em um barramento levam a:
 - Atrasos de propagação
 - Longos caminhos de dados significa que a coordenação do uso do barramento pode afetar contrariamente o desempenho.
 - Se a demanda de transferência de dados agregada se aproxima da capacidade do barramento.
- A maioria dos sistemas utiliza múltiplos barramentos para contornar esses problemas.

Estrutura de barramento tradicional (ISA) (com cache)



Arquitetura de alto desempenho



Tipos de barramento

- Dedicado:
 - Linhas separadas para dados e endereço.
- Multiplexado.
 - Linhas compartilhadas.
 - Linha de controle válida de endereço ou dados.
 - Vantagem
 - Menos linhas
 - Desvantagens:
 - Controle mais complexo.
 - Desempenho máximo.

Arbitração de barramento

- Mais de um módulo controlando o barramento.
- P.e., CPU e controlador de DMA.
- Apenas um módulo pode controlar barramento de uma só vez.
- Arbitração pode ser centralizada ou distribuída.

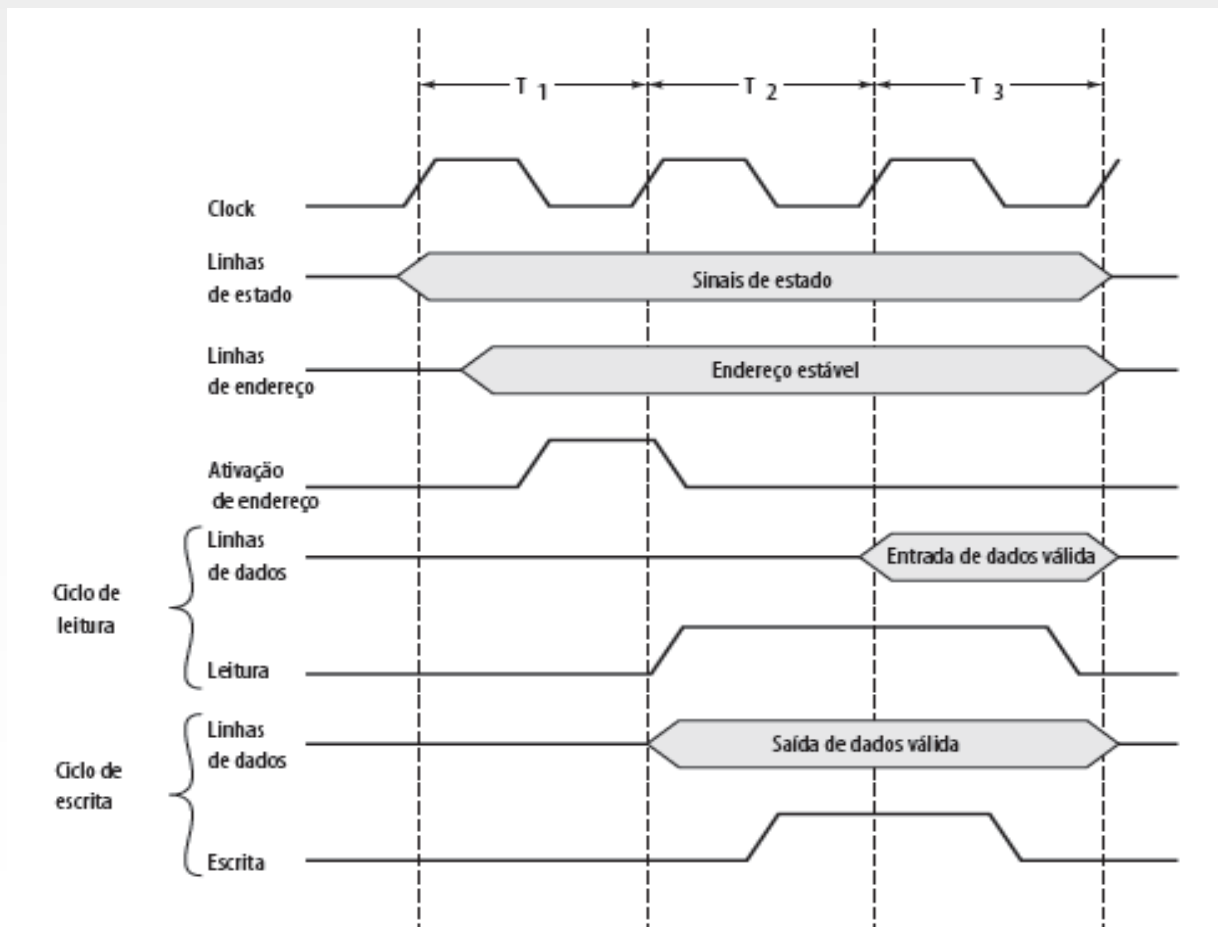
Arbitração centralizada e distribuída

- Centralizada:
 - Único dispositivo de hardware controlando o acesso ao barramento.
 - Controlador de barramento.
 - Árbitro.
 - Pode ser parte da CPU ou separada.
- Distribuída:
 - Cada módulo pode reivindicar o barramento.
 - Lógica de controle em todos os módulos.

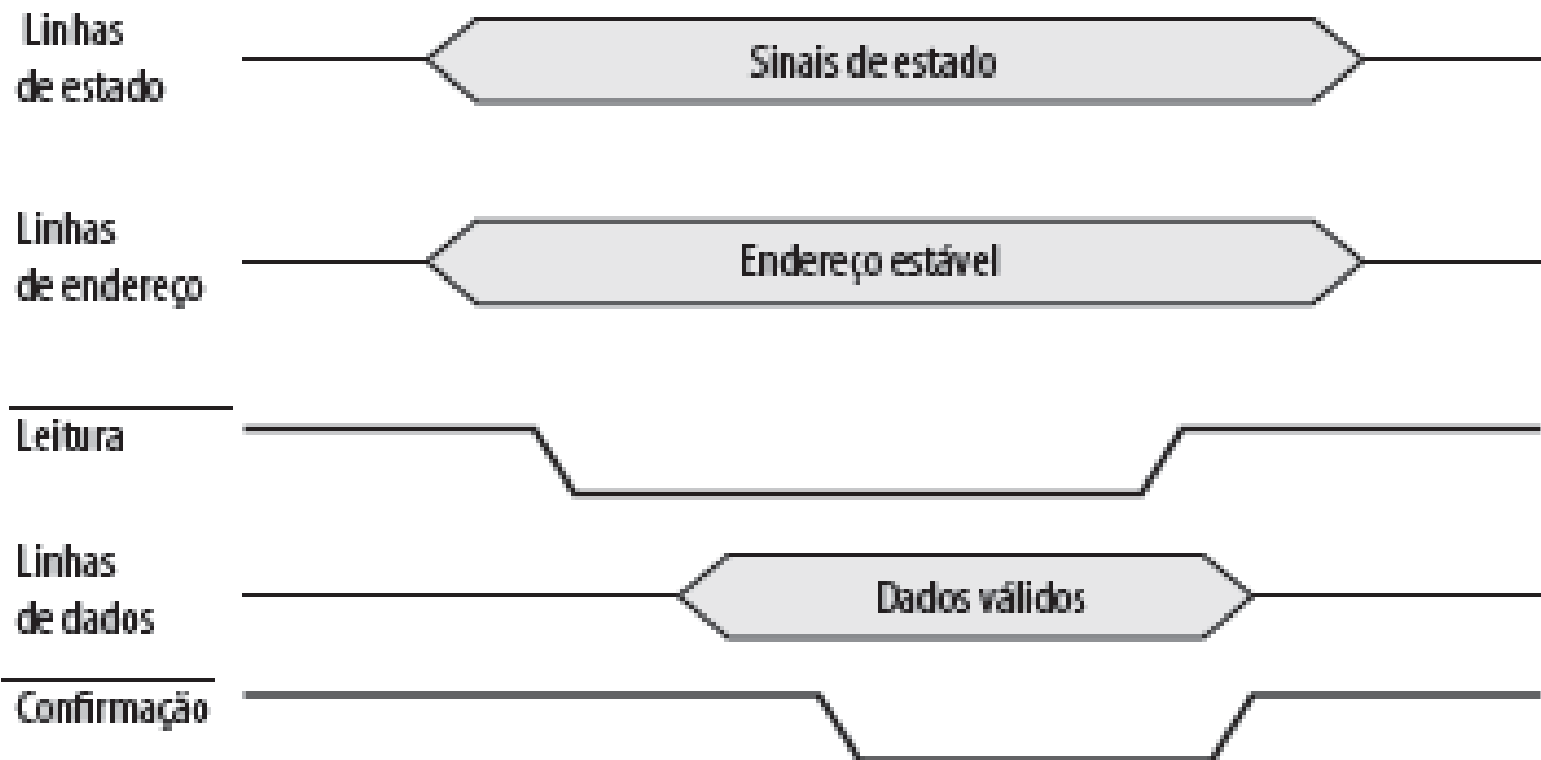
Temporização

- Coordenação de eventos no barramento.
- Síncrona:
 - Eventos determinados por sinais de clock.
 - Barramento de controle inclui linha de clock.
 - Uma única transmissão 1-0 é um ciclo do barramento.
 - Todos os dispositivos podem ler linha de clock.
 - Normalmente, sincronismo na borda inicial.
 - Geralmente, um único ciclo para um evento.

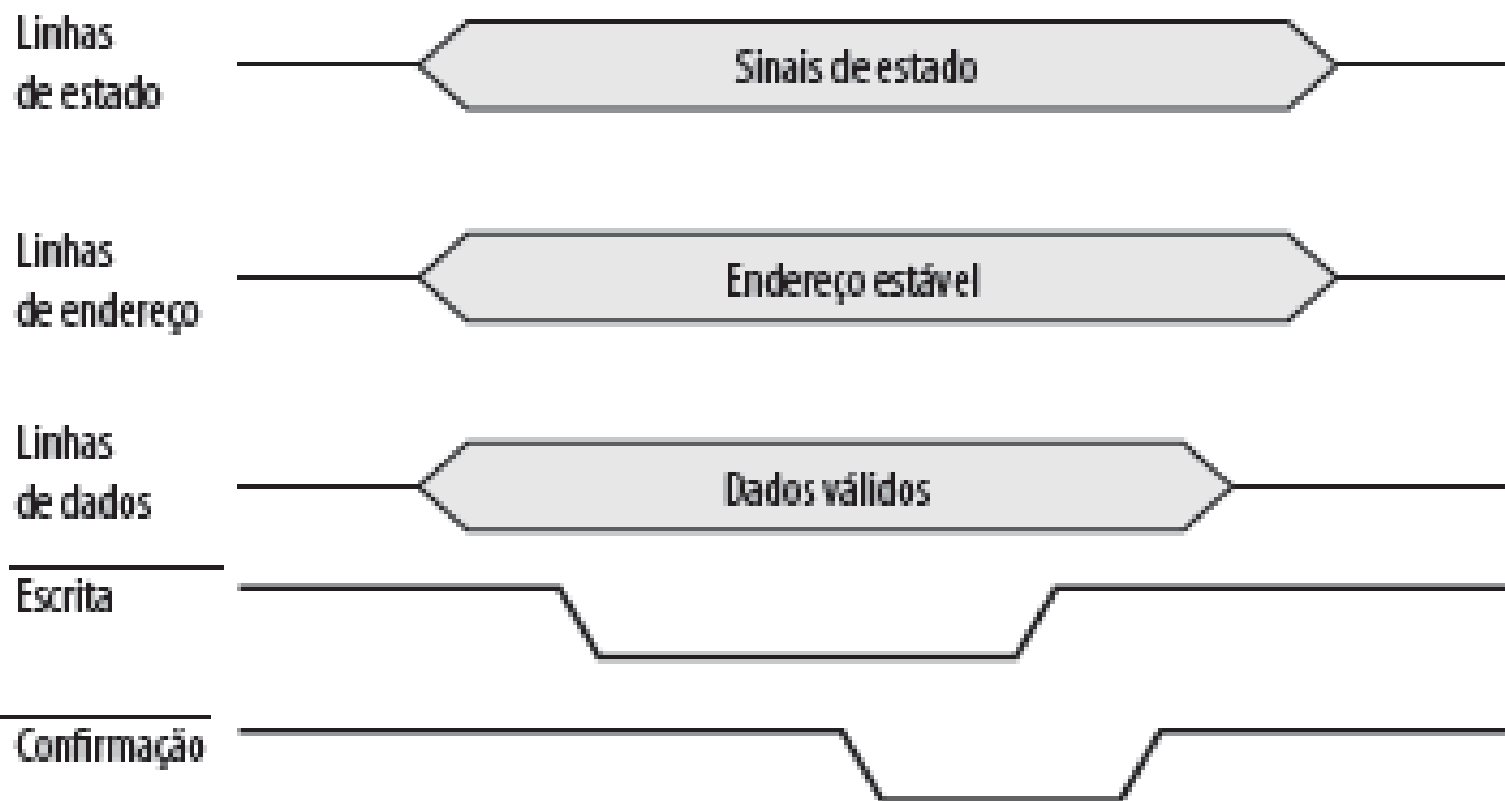
Diagrama de temporização síncrona



Temporização assíncrona – diagrama de leitura



Temporização assíncrona – diagrama de escrita



Barramento PCI

- Interconexão de componente periférico.
- Intel lançado para domínio público.
- 32 ou 64 bits.
- 50 linhas.

Linhas de barramento PCI (obrigatórias)

- Linhas de sistemas:
 - Incluindo clock e reset.
- Endereços e dados:
 - 32 linhas multiplexadas para endereços e dados.
 - Linhas de interrupção e validação.
 - Controle da interface.
- Arbitração:
 - Não compartilhada.
 - Conexão direta com barramento PCI.
 - Linhas de erro.

Comandos PCI

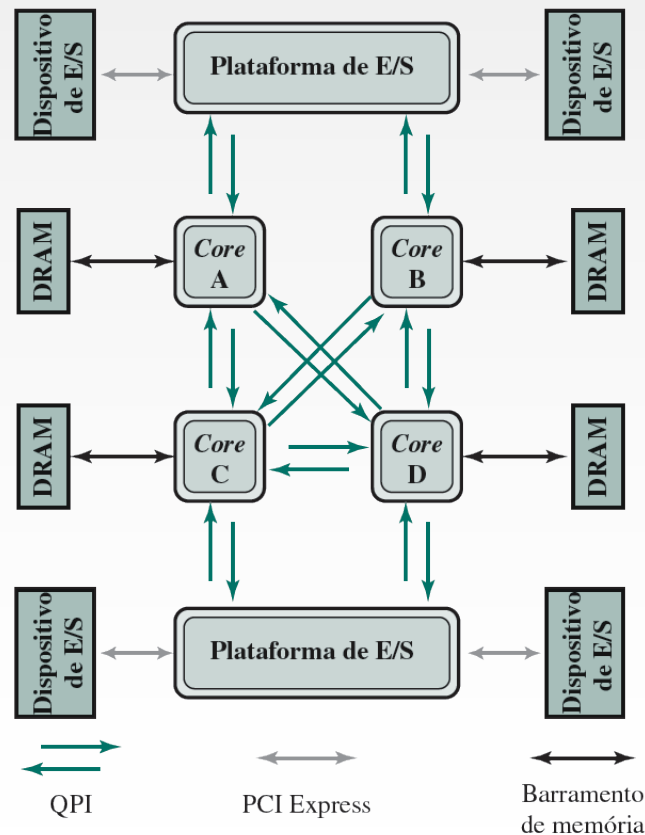
- Transação entre iniciador (mestre) e destino.
- Mestre reivindica barramento.
- Determina tipo de transação.
- P.e., leitura/escrita de E/S.
- Fase de endereço.
- Uma ou mais fases de dados.

Interconexão ponto a ponto

- Em comparação com o barramento compartilhado, a **interconexão ponto a ponto** tem menor latência, maior taxa de dados e melhor escalabilidade.
- Analisamos um exemplo importante e representativo da técnica de interconexão ponto a ponto: **QuickPath Interconnect (QPI)**, da Intel, que foi apresentada em 2008.
- Características significativas da QPI:
 - **Conexões diretas múltiplas**
 - **Arquitetura de protocolo em camadas**
 - **Transferência de dados em pacotes**

Interconexão ponto a ponto

- Configuração multicore usando QPI:



Interconexão ponto a ponto

- Define-se QPI como uma arquitetura de protocolo de quatro camadas, que abrange as seguintes camadas:

1. Física

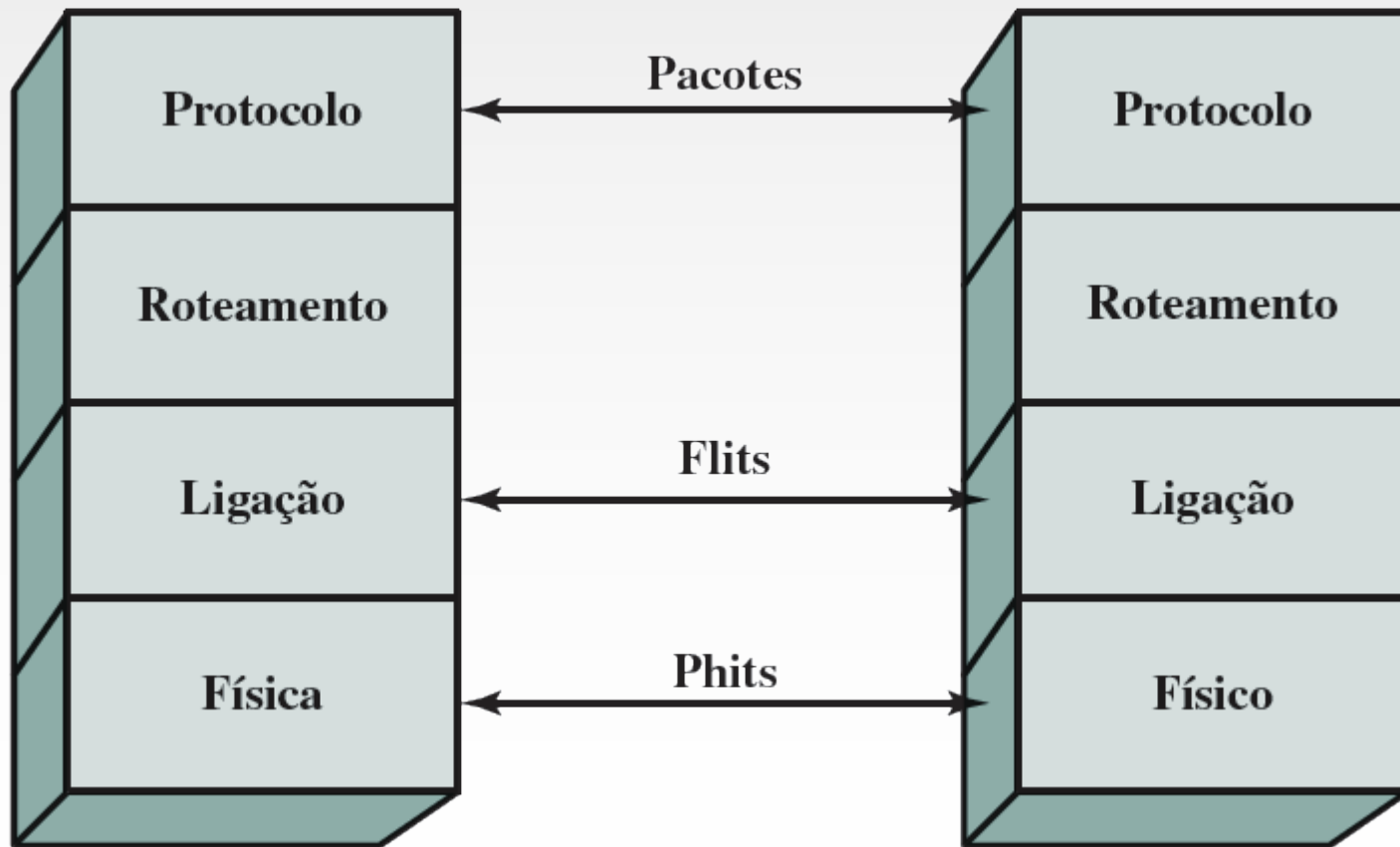
1. Ligação

1. Roteamento

1. Protocolo

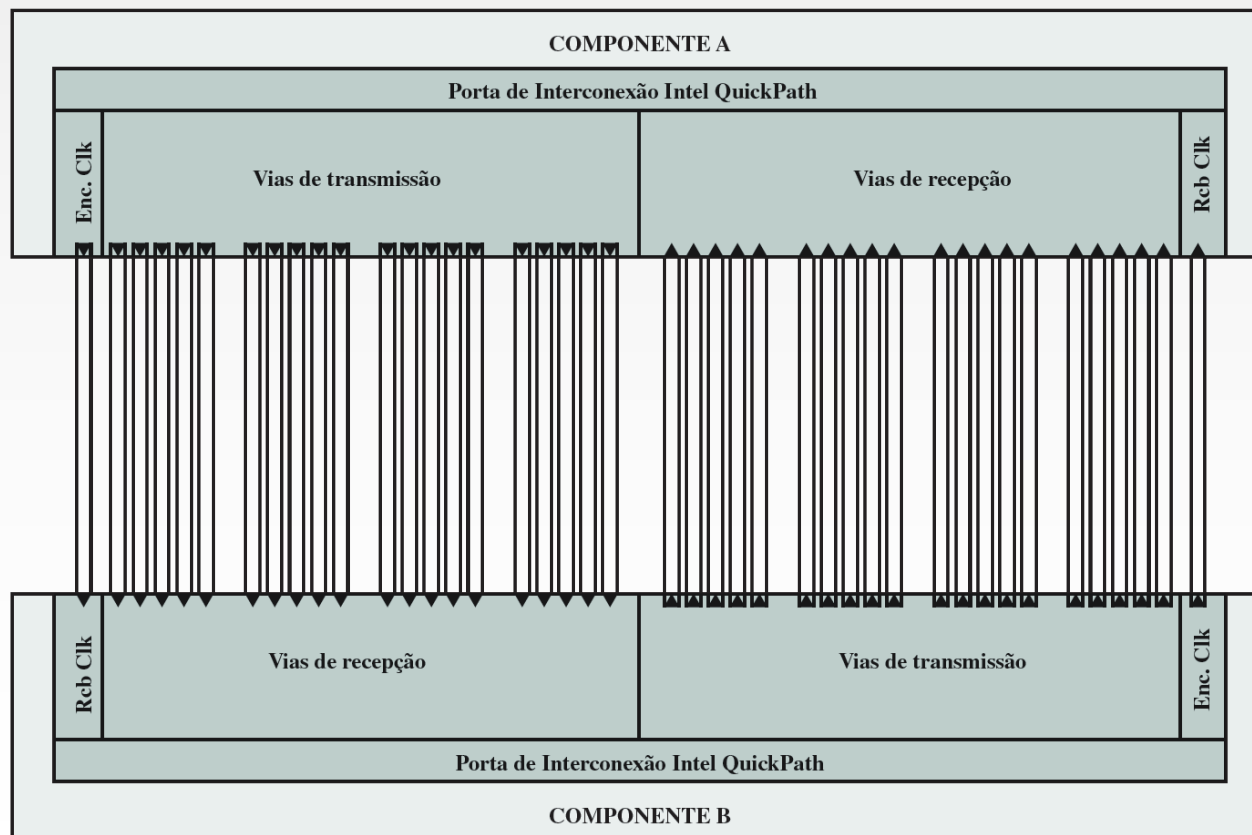
- A figura a seguir mostra camadas de QPI.

Interconexão ponto a ponto



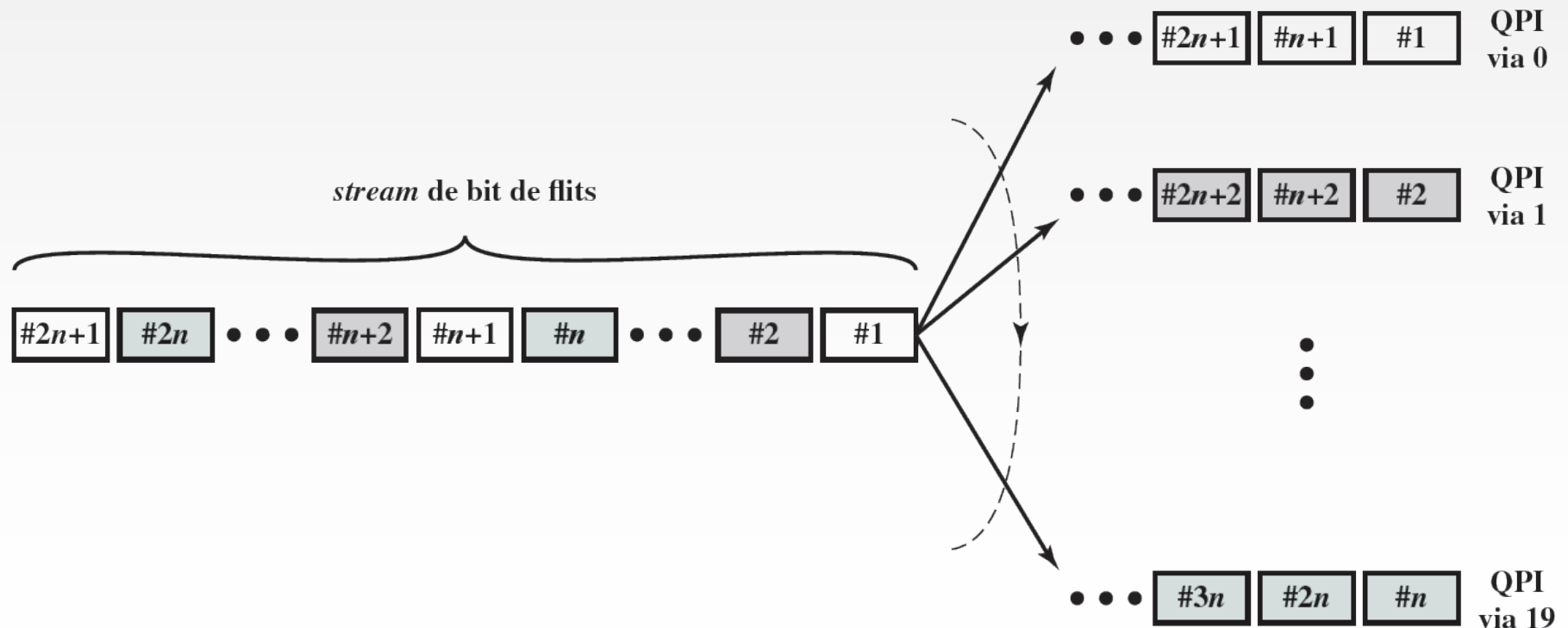
QPI – Camada física

- Interface física da interconexão da Intel QPI:



QPI – Camada física

➤ Distribuição multivia da QPI:



QPI – Camada de ligação

- A **camada de ligação** QPI desempenha duas funções-chave: controle de fluxo e controle de erro.
- A **função de controle de fluxo** é necessária para assegurar que a entidade de QPI de envio não sobrecarregue uma entidade de QPI de recebimento ao enviar dados mais rápido do que o receptor pode processar e para limpar os buffers para receberem mais dados.
- A **função de controle de erro** em uma camada de ligação detecta e recupera a partir desses erros de bits, e então isola camadas mais altas a partir da experiência de erros de bits.

QPI – Camada de roteamento

- A **camada de roteamento** é usada para determinar o caminho que um pacote vai trafegar através de interconexões disponíveis do sistema.
- Em pequenas configurações as opções de roteamento são limitadas e as tabelas de roteamento são bastante simples.
- Para sistemas maiores, as opções de tabela de roteamento são mais complexas, dependendo de como (1) dispositivos são alocados na plataforma, (2) recursos do sistema são divididos e (3) eventos de confiabilidade resultam no mapeamento em torno de um recurso de falha.

QPI – Camada de protocolo

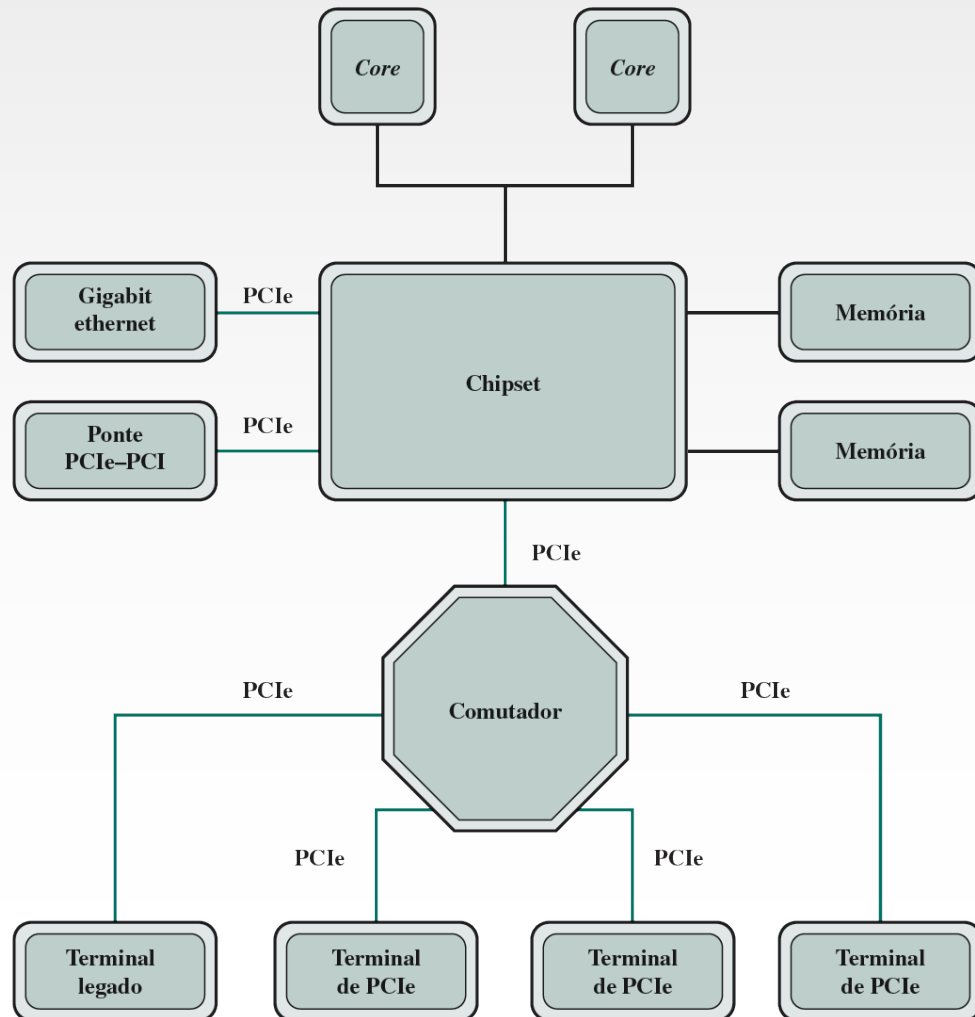
- Na camada, o pacote é definido como uma unidade de transferência.
- A definição de conteúdo de pacote é padronizada com alguma flexibilidade permitida ao atender pedidos diferentes de segmentos de mercado.
- Uma função-chave desempenhada é um protocolo de coerência de cache, que age se certificando de que os valores da memória principal mantidos em diversas caches são consistentes.
- Uma carga útil de pacote de dados comum é um bloco de dados enviados para e a partir de uma cache.

PCI Express

- Após anos de uso do barramento PCI, uma nova versão do barramento, conhecida como **PCI Express (PCIe)** foi desenvolvida.
- A PCIe é um esquema de interconexão ponto a ponto que visa substituir os esquemas baseados em barramento, como a PCI.

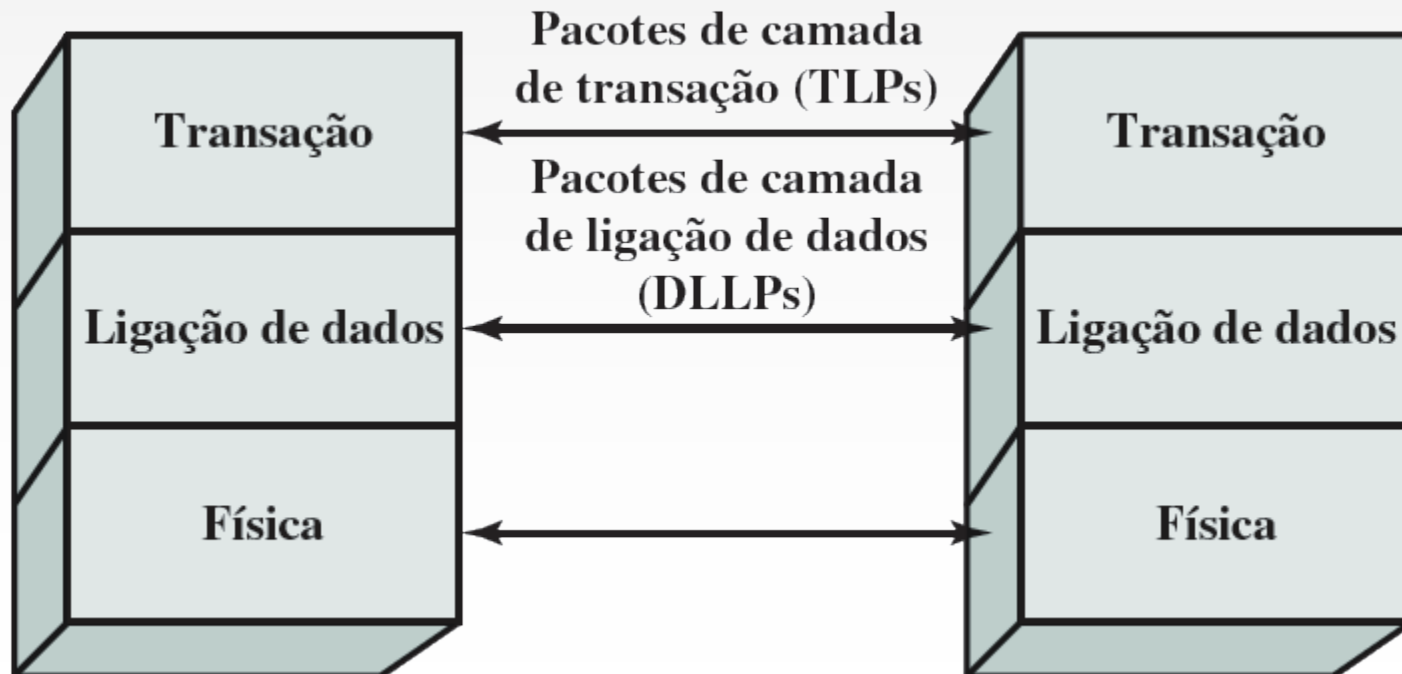
PCI – Arquitetura física e lógica

Configuração comum usando PCIe:



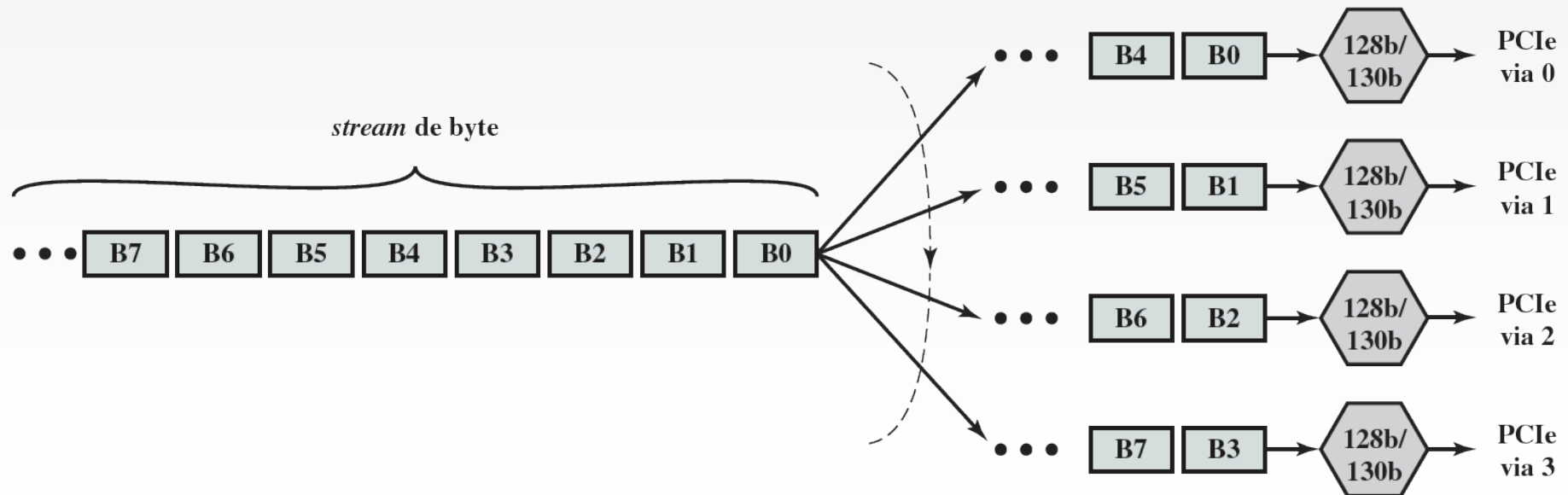
PCI – Arquitetura física e lógica

- Camadas de protocolo de PCIe:



PCIe – Camada física

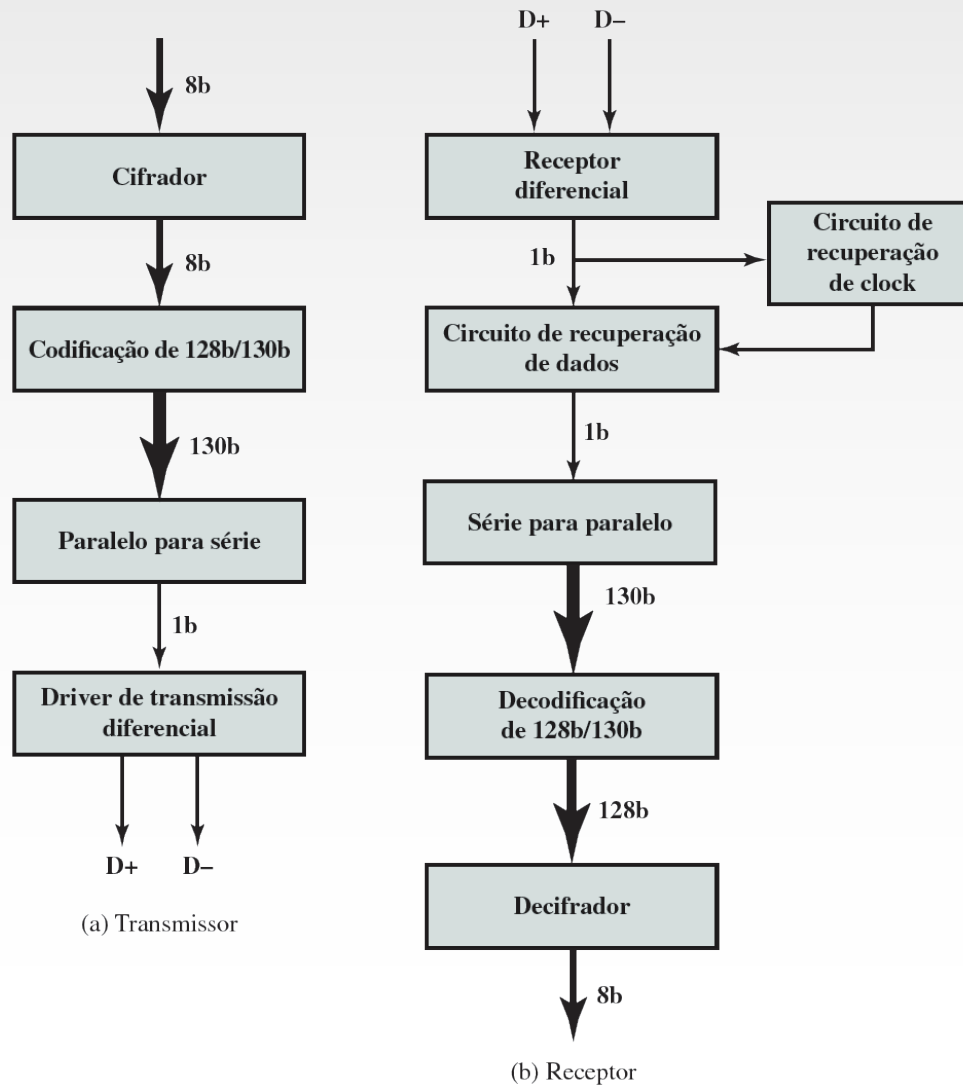
- Técnica de distribuição multivia do PCIe:



PCIe – Camada física

- Uma técnica comum, e a única usada na PCIe 3.0, para superar o problema de uma longa cadeia de bits de um valor é a cifragem.
- A **cifragem** é uma técnica de mapeamento que tende a fazer os dados aparecerem de modo mais aleatório.
- Outra técnica que pode auxiliar na sincronização é a **codificação**, em que bits adicionais são inseridos no stream de bits para forçar transições.
- A figura a seguir ilustra o uso da cifragem e da codificação.

PCIe – Camada física



- A **camada de transação (TL)** recebe pedidos de leitura e escrita a partir do software acima da TL e cria pacotes de solicitação de transmissão para um destino por meio da camada de ligação.
- A TL suporta quatro espaços endereçados:
 - 1. Memória**
 - 1. E/S**
 - 1. Configuração**
 - 1. Mensagem**

Tipos de transação PCIe TLP

Espaço de endereço	Tipo de TLP	Propósito
Memória	Pedido de leitura de memória	Transfere os dados para ou a partir de um local no mapa de memória do sistema
	Pedido de bloqueio de leitura de memória	
	Pedido de escrita de memória	
E/S	Pedido de leitura de E/S	Transfere os dados para ou a partir de um local no mapa de memória do sistema para dispositivos legado
	Pedido de escrita de E/S	
Configuração	Pedido de leitura de tipo de configuração 0	Transfere os dados para e a partir de um local no espaço de configuração de um dispositivo PCIe
	Pedido de escrita de tipo de configuração 0	
	Pedido de leitura de tipo de configuração 1	
	Pedido de escrita de tipo de configuração 1	
Mensagem	Pedido de mensagem	Proporciona uma mensagem em banda e um relato de evento
	Pedido de mensagem com dados	
Memória, E/S, Configuração	Finalização	Retornado para certos pedidos
	Finalização com dados	
	Finalização bloqueada	
	Finalização bloqueada para dados	

Referências:

- William Stallings, Arquitetura e Organização de Computadores, 8ª ed. 2010
- Mario A. Monteiro, Introdução à Organização de Computadores, 5ª ed. 2007