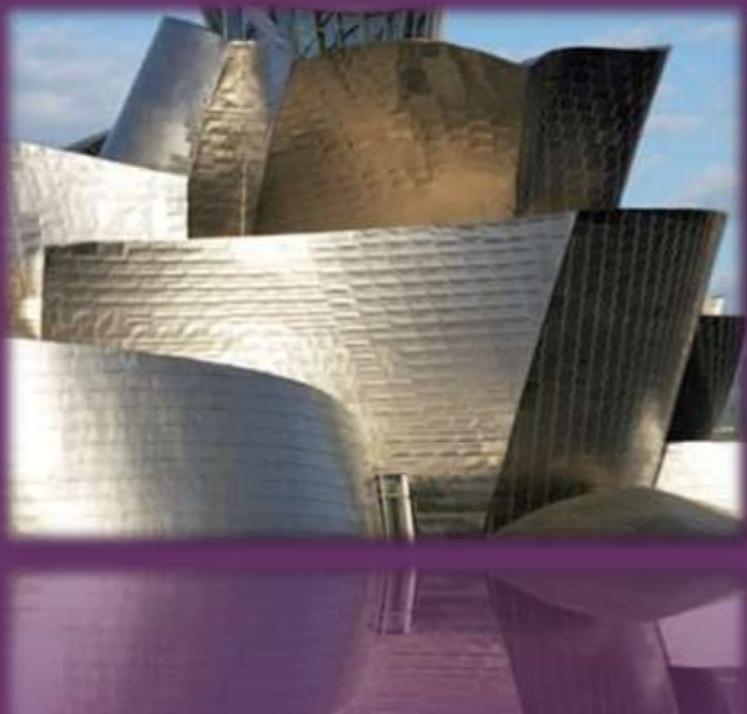


+



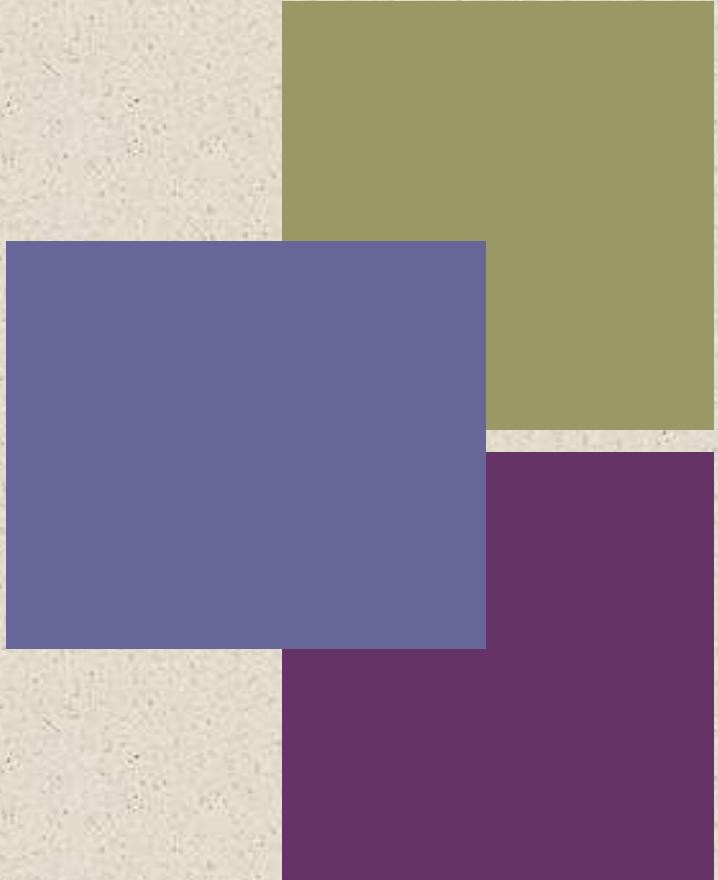
Arquitetura de Computadores

Prof. Marcial Fernández

2025.2

+ Capítulo 7

Entrada/Saída



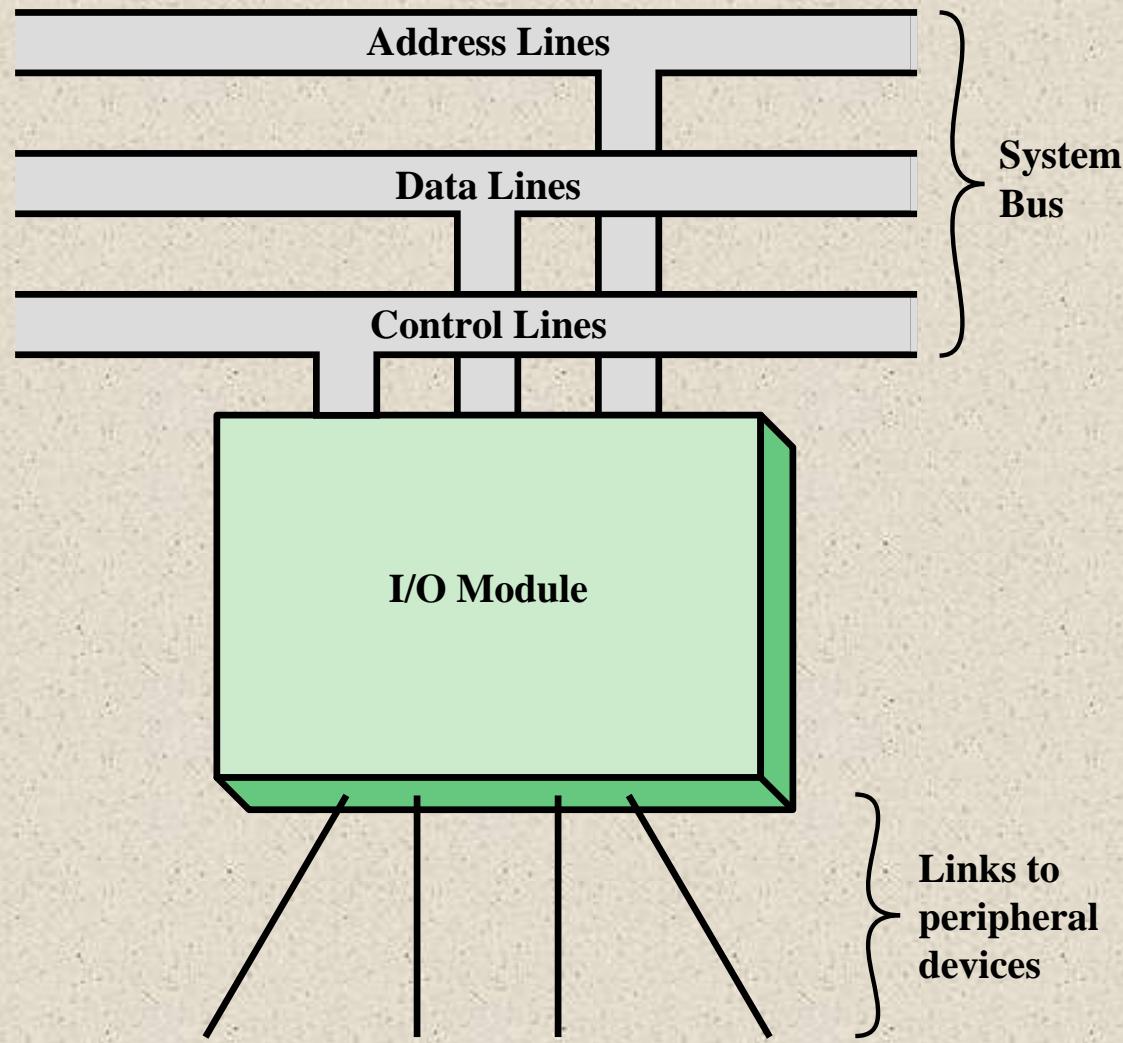


Figure 7.1 Generic Model of an I/O Module

+

Dispositivos externos



- Fornecer um meio de troca de dados entre o ambiente externo e o computador
- Conectar o computador por meio de um link com módulos de E/S
 - O link é usado para trocar controle, status e dados entre o módulo de E/S e o dispositivo externo
- Dispositivo periférico
 - Um dispositivo externo conectado a um módulo de E/S

Três categorias:

- Legível por humanos
 - Adequado para comunicação com o usuário do computador
 - Terminal de vídeo, impressoras
- Legível por máquina
 - Adequado para comunicação com equipamentos
 - Sistemas de disco externo, sensores e atuadores
- Comunicação
 - Adequado para comunicação com dispositivos remotos, como um terminal, ou outro computador

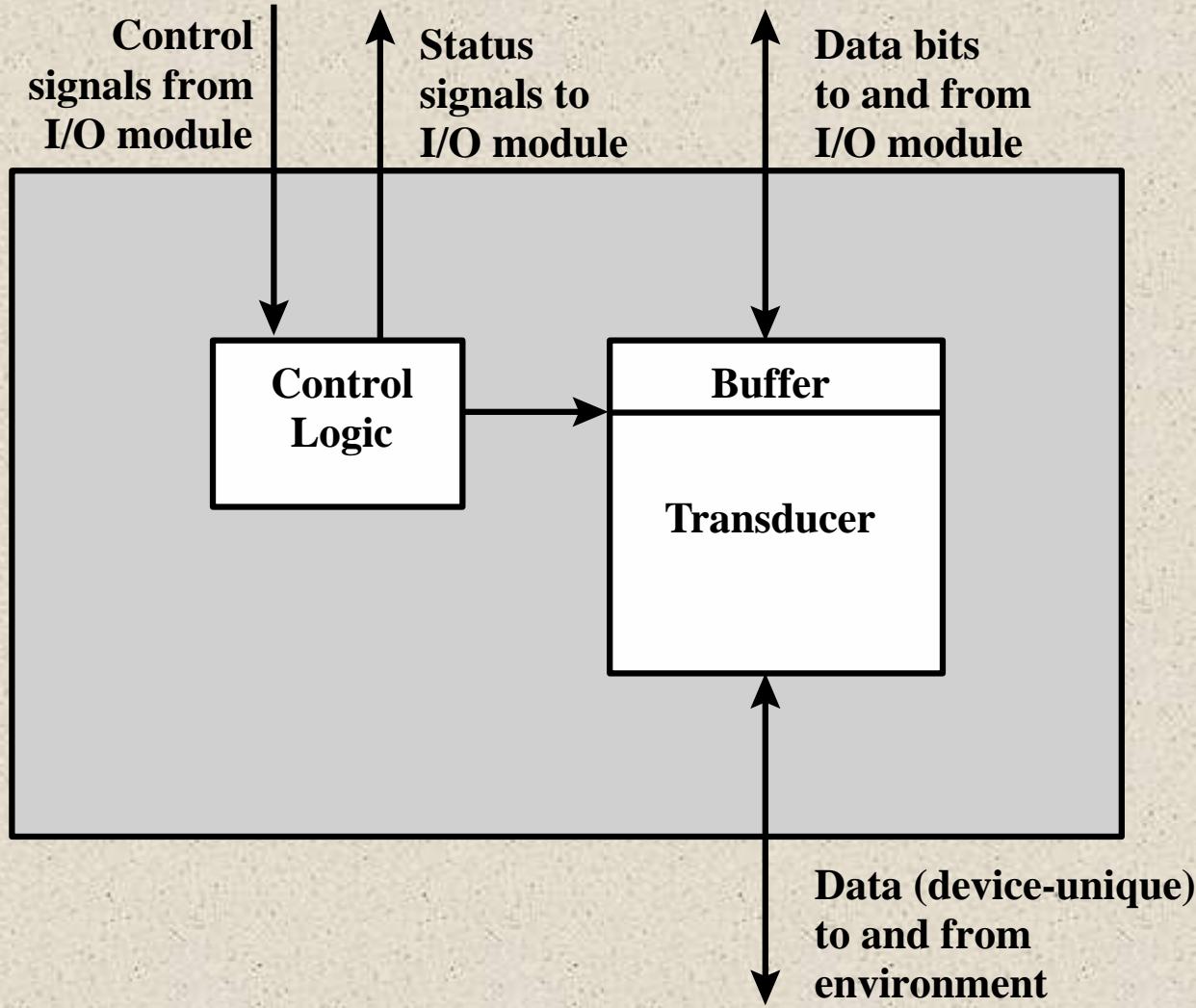


Figure 7.2 Block Diagram of an External Device



Teclado/Monitor

International Reference Alphabet (IRA)

- A unidade básica de troca é o caractere
 - Um código é associado a cada caractere
 - Cada caractere é representado por um código binário exclusivo de 7 bits
 - Podem ser representados 128 caracteres diferentes
- Os caracteres são de dois tipos:
 - Imprimível
 - Caracteres alfabéticos, numéricos e especiais que podem ser impressos em papel ou exibidos em uma tela
 - Controle
 - Tem a ver com o controle da impressão ou exibição de caracteres
 - Exemplo é retorno de carro (CR)
 - Outros caracteres de controle estão relacionados aos procedimentos de comunicação

Meios mais comuns de interação computador/usuário

- usuário fornece a entrada por meio do teclado
- monitor exibe dados fornecidos pelo computador

Códigos de teclado

- Quando o usuário pressiona uma tecla, é gerado um sinal eletrônico que é interpretado pelo transdutor no teclado e traduzido no padrão de bits do código IRA correspondente.
- Este padrão de bits é transmitido para o módulo de E/S no computador
- Na saída, os caracteres do código IRA são transmitidos para um dispositivo externo do módulo de E/S
- O transdutor interpreta o código e envia os sinais eletrônicos necessários ao dispositivo de saída para exibir o caractere indicado ou executar a função de controle solicitada

As principais funções de um módulo de E/S se enquadram nas seguintes categorias:

Controle e tempo

- Coordena o fluxo de tráfego entre recursos internos e dispositivos externos

Comunicação do processador

- Envolve decodificação de comando, dados, relatórios de status, reconhecimento de endereço

Comunicação do dispositivo

- Envolve comandos, informações de status e dados

Buffering de dados

- Executa a operação de buffer necessária para equilibrar as velocidades do dispositivo e da memória

Detecção de erro

- Detecta e relata erros de transmissão

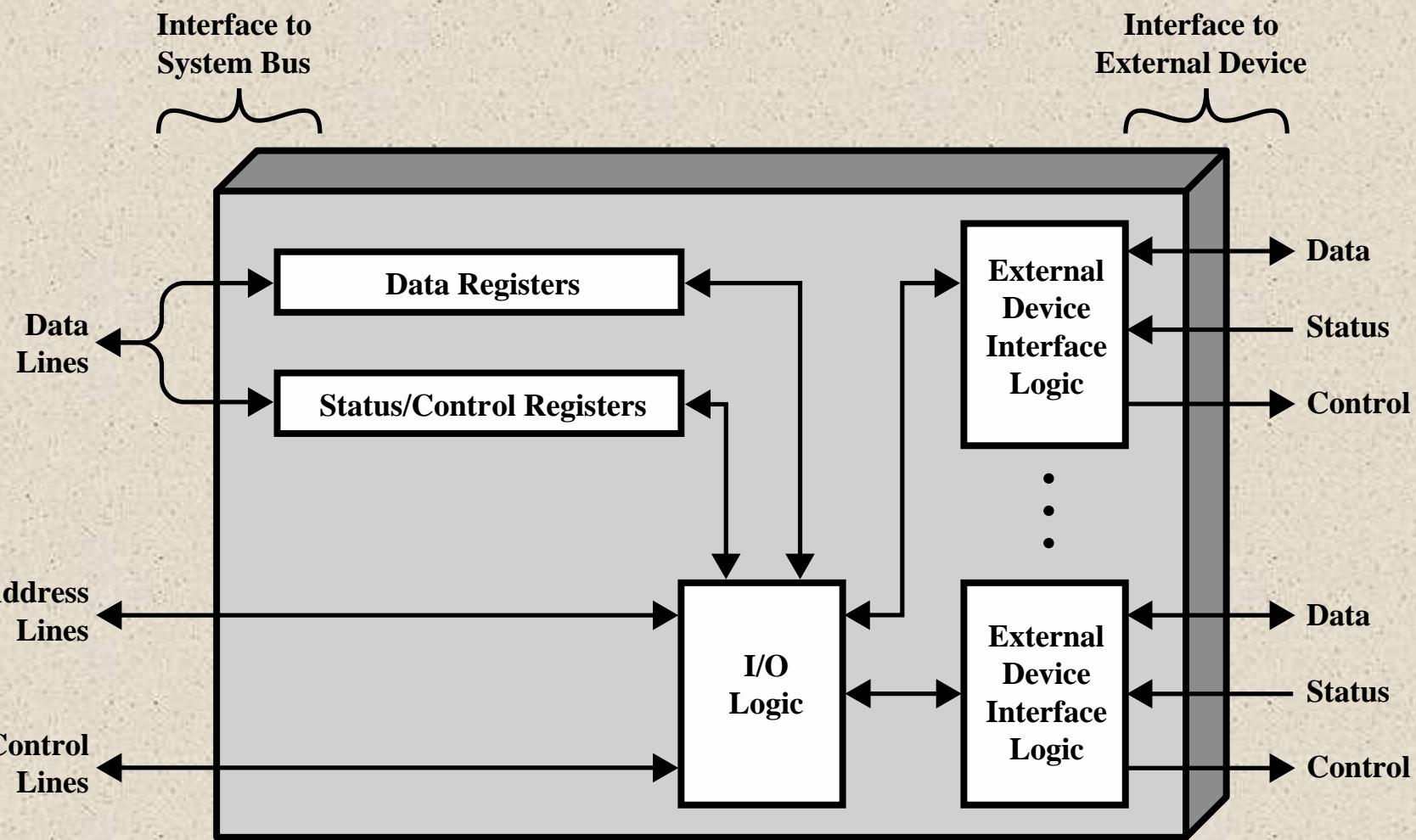


Figure 7.3 Block Diagram of an I/O Module

+ E/S programada

Três técnicas são possíveis para operações de E/S:

- E/S programada
 - Os dados são trocados entre o processador e o módulo de E/S a partir de iniciativa do processador
 - O processador executa um programa que lhe dá controle direto da operação de E/S
 - Quando o processador emite um comando, ele deve esperar até que a operação de E/S seja concluída
 - Se o processador for mais rápido que o módulo de E/S, provoca desperdício de tempo no processador
- E/S acionada por interrupção
 - O processador emite um comando de E/S, continua a executar outras instruções e é interrompido pelo módulo de E/S quando este conclui seu trabalho
- Acesso direto à memória (DMA)
 - O módulo de E/S e a memória principal trocam dados diretamente sem envolvimento do processador

Tabela 7.1

Técnicas de E/S

	No Interrupts	Use of Interrupts
I/O-to-memory transfer through processor	Programmed I/O	Interrupt-driven I/O
Direct I/O-to-memory transfer		Direct memory access (DMA)

Comandos de E/S

- Existem quatro tipos de comandos de E/S que um módulo de E/S pode receber quando é endereçado por um processador:

1) Controlar

- usado para ativar um periférico e dizer a ele o que fazer

2) Teste

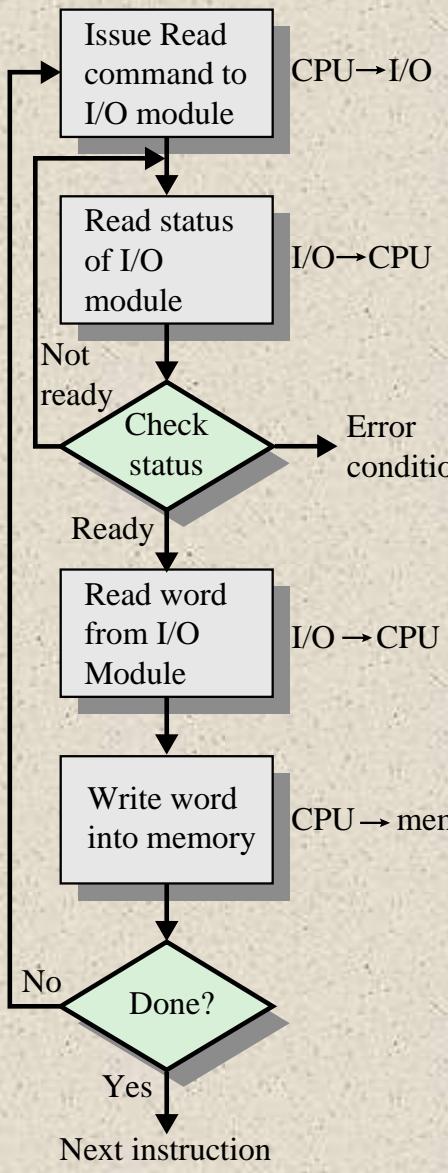
- usado para testar várias condições de status associadas a um módulo de E/S e seus periféricos

3) Ler

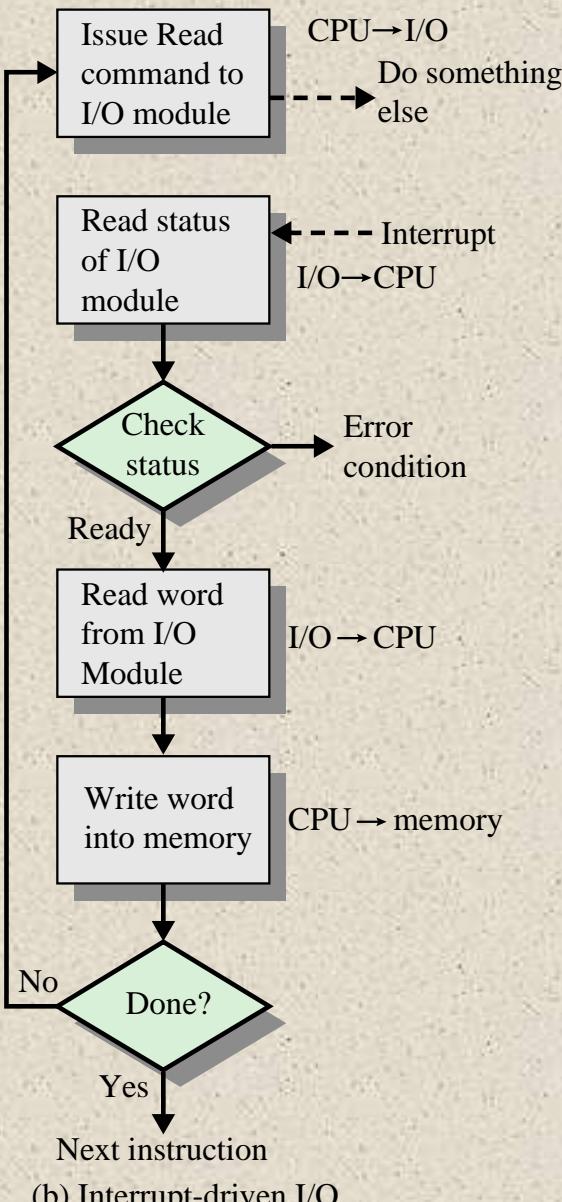
- faz com que o módulo de E/S obtenha dados do periférico e o coloque em um buffer interno

4) Escrever

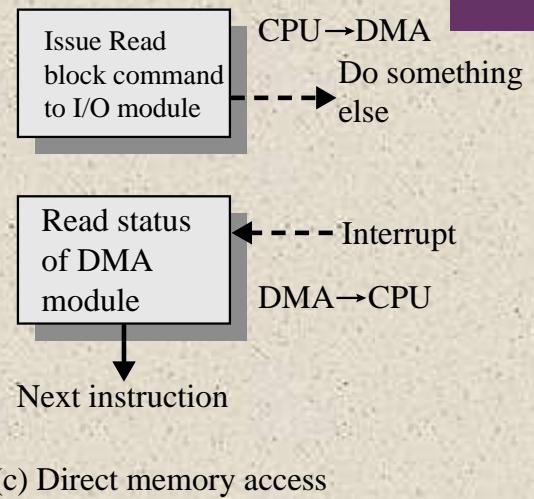
- faz com que o módulo de E/S pegue dados do barramento de dados e posteriormente transmita esses dados para o periférico



(a) Programmed I/O



(b) Interrupt-driven I/O



(c) Direct memory access

Figure 7.4 Three Techniques for Input of a Block of Data

Instruções de E/S

Com E/S programada, há uma correspondência próxima entre as instruções relacionadas à E/S que o processador busca na memória e os comandos de E/S que o processador emite para um módulo de E/S para executar as instruções.

A forma da instrução depende da maneira como os dispositivos externos são endereçados

Cada dispositivo de E/S conectado por meio de módulos de E/S recebe um identificador ou endereço exclusivo

Quando o processador emite um comando de E/S, o comando contém o endereço do dispositivo desejado

Assim, cada módulo de E/S deve interpretar as linhas de endereço para determinar se o comando é para si mesmo

E/S mapeadas em memória

Existe um único espaço de endereço para locais de memória e dispositivos de E/S

Uma única linha de leitura e uma única linha de gravação são necessárias no barramento

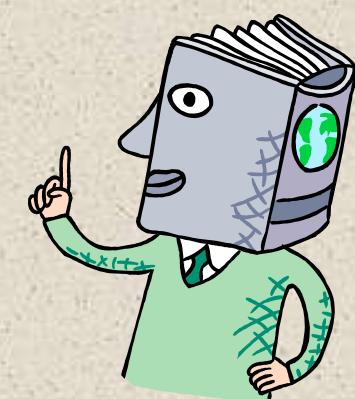
Resumo do mapeamento de E/S

■ E/S mapeada em memória

- Dispositivos e memória compartilham um espaço de endereço
- E/S se parece com leitura/gravação de memória
- Nenhum comando especial para E/S
 - Usa o conjunto de comandos de acesso à memória disponíveis

■ E/S isolada

- Espaço de endereço separados
- Precisa de linhas de seleção de E/S ou memória
- Comandos especiais para E/S
 - Conjunto de comandos limitado



E/S acionada por interrupção

O problema com E/S programada é que o processador tem que esperar muito tempo para que o módulo de E/S esteja pronto para receber ou transmitir dados

Uma alternativa é o processador emitir um comando de E/S para um módulo e então prosseguir com algum outro trabalho

O módulo de E/S interromperá o processador para solicitar serviço quando estiver pronto para trocar dados com o processador

O processador executa a transferência de dados e retoma seu processamento anterior

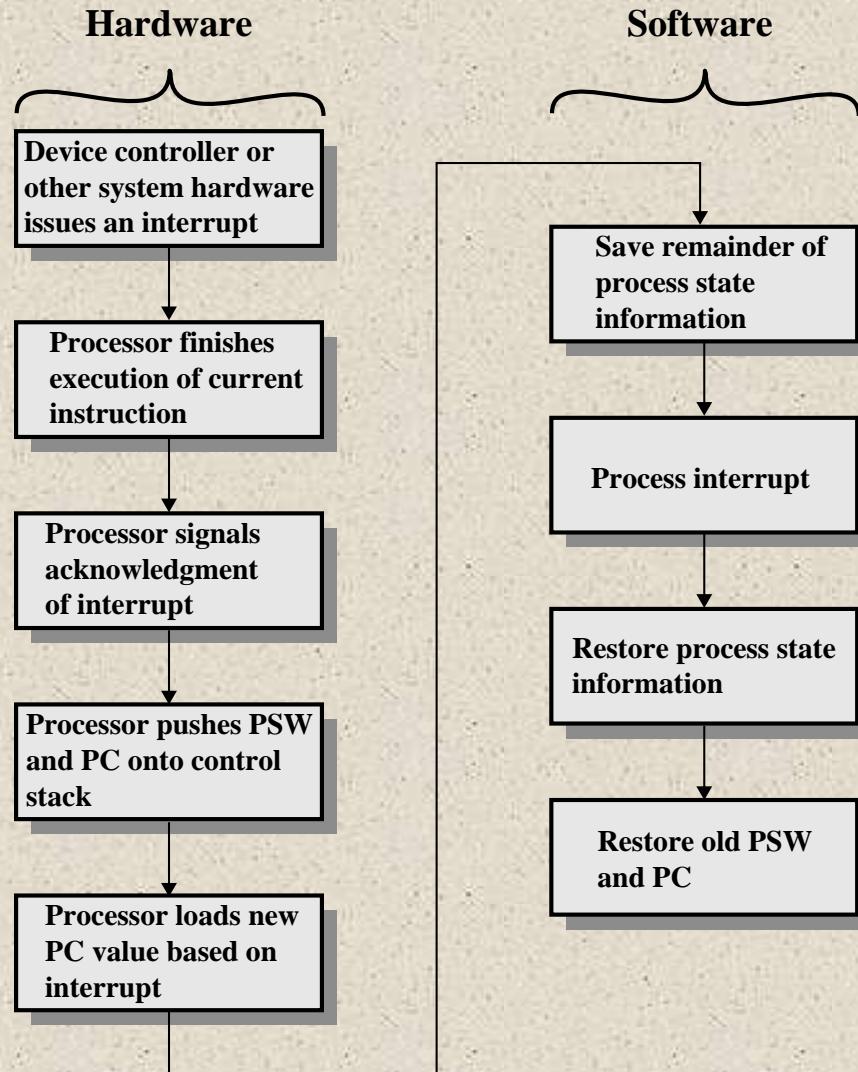


Figure 7.6 Simple Interrupt Processing

Problemas de design

Dois problemas de design surgem na implementação de E/S de interrupção:

- Como haverá vários módulos de E/S, como o processador determina qual dispositivo emitiu a interrupção?
- Se várias interrupções ocorrerem, como o processador decide qual delas processar?



Identificação do dispositivo

- **Várias linhas de interrupção**
 - Entre o processador e os módulos de E/S
 - Abordagem mais direta para o problema
 - Consequentemente, mesmo que várias linhas sejam usadas, é provável que cada linha tenha vários módulos de E/S conectados a ela
- **Enquete de software**
 - Quando o processador detecta uma interrupção, ele ramifica para uma rotina de serviço de interrupção cuja tarefa é consultar cada módulo de E/S para determinar qual módulo solicitou a interrupção.
 - Demorado
- **Daisy chain (enquete de hardware, vetorizada)**
 - A linha de reconhecimento de interrupção é encadeada em série pelos módulos
 - Vetor – endereço do módulo de E/S ou algum outro identificador exclusivo
 - Interrupção vetorizada – o processador usa o vetor como um ponteiro para a rotina de serviço de dispositivo apropriada, evitando a necessidade de executar primeiro uma rotina geral de serviço de interrupção
- **Arbitragem de barramento (vetorizada)**
 - Um módulo de E/S deve primeiro obter o controle do barramento antes de poder gerar a linha de solicitação de interrupção
 - Quando o processador detecta a interrupção, ele responde na linha de reconhecimento de interrupção
 - Em seguida, o módulo solicitante coloca seu vetor nas linhas de dados

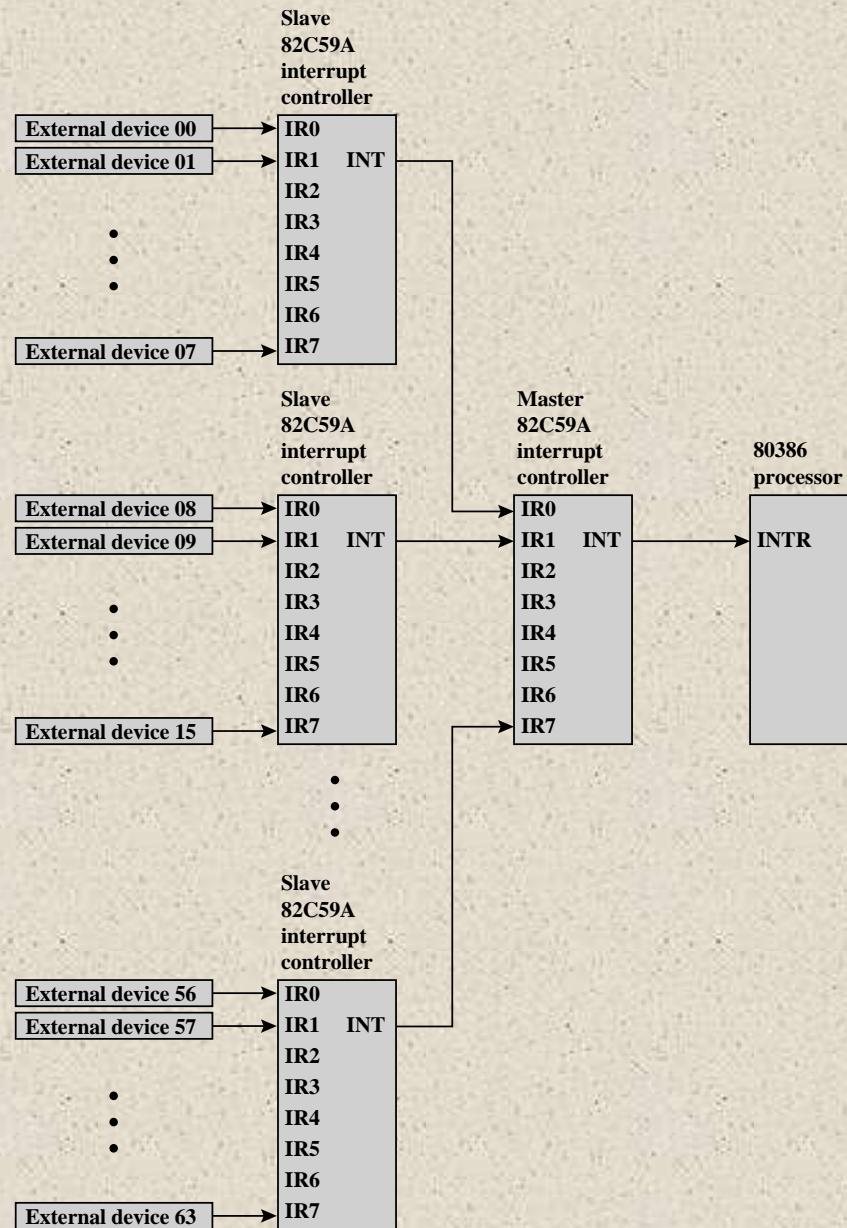
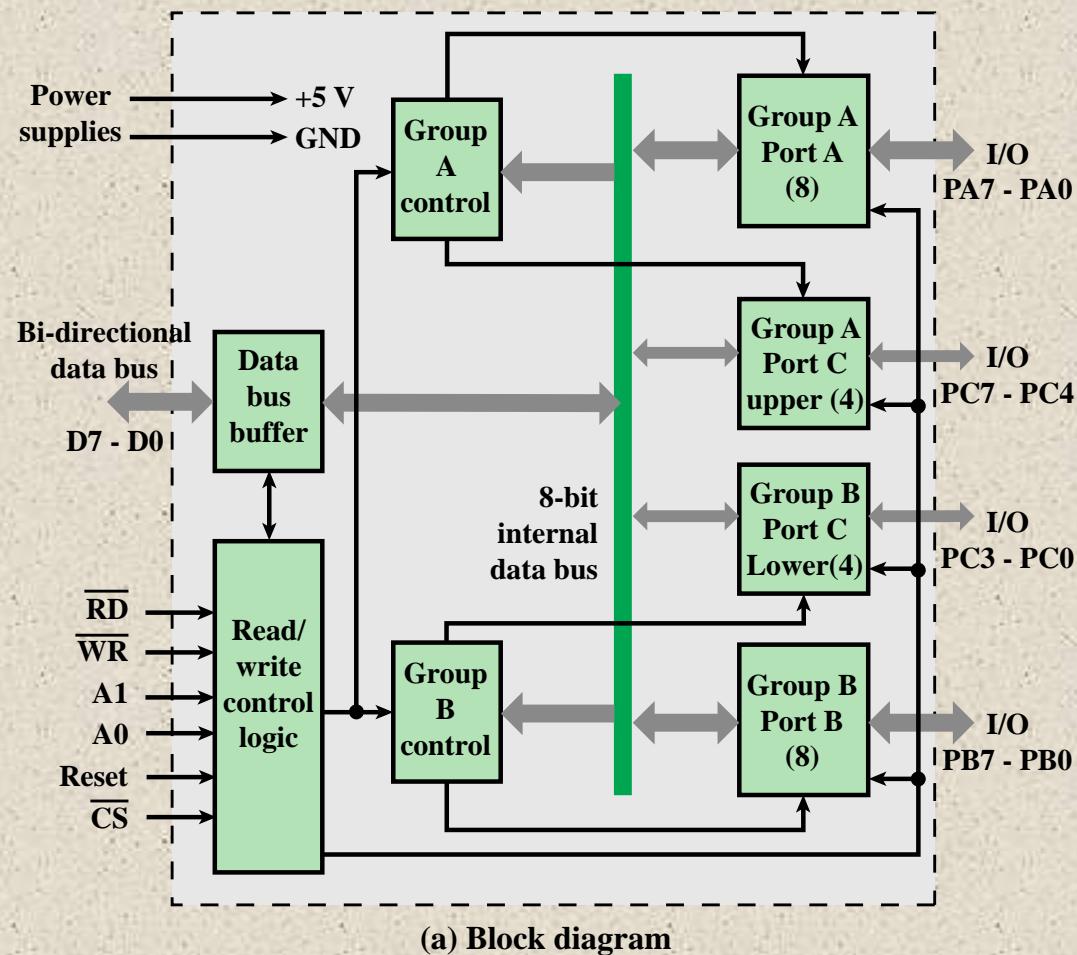


Figure 7.8 Use of the 82C59A Interrupt Controller



(a) Block diagram

PA3	1	40	PA4
PA2	2	39	PA5
PA1	3	38	PA6
PA0	4	37	PA7
RD	5	36	WR
CS	6	35	Reset
GND	7	34	D0
A1	8	33	D1
A0	9	8255A	32
PC7	10	31	D3
PC6	11	30	D4
PC5	12	29	D5
PC4	13	28	D6
PC3	14	27	D7
PC2	15	26	V
PC1	16	25	PB7
PC0	17	24	PB6
PB0	18	23	PB5
PB1	19	22	PB4
PB2	20	21	PB3

(b) Pin layout

Figure 7.9 The Intel 8255A Programmable Peripheral Interface

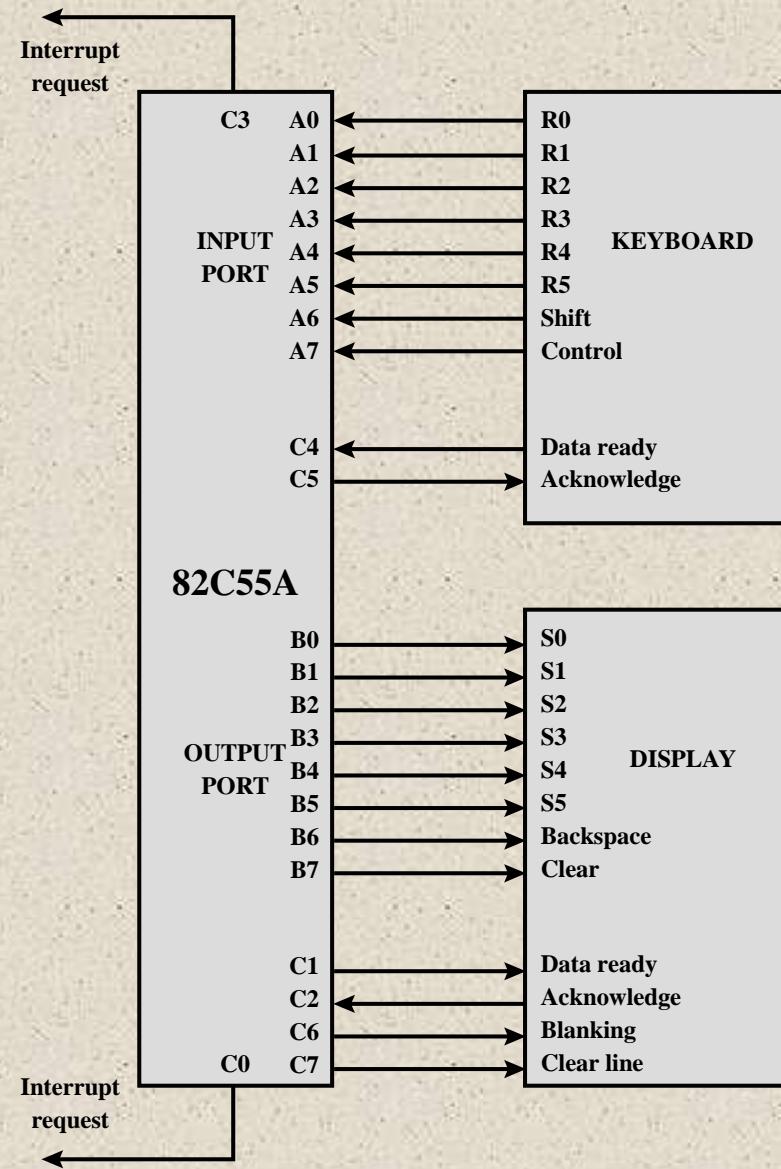


Figure 7.11 Keyboard/Display Interface to 82C55A

Desvantagens da E/S programada e controlada por interrupção

- Ambas as formas de E/S sofrem de duas desvantagens inerentes:
 - 1) A taxa de transferência de E/S é limitada pela velocidade com que o processador pode testar e servir um dispositivo
 - 2) O processador está vinculado ao gerenciamento de uma transferência de E/S; um número de instruções deve ser executada para cada transferência de E/S
- Quando grandes volumes de dados precisam ser movidos, a técnica mais eficiente é o acesso direto à memória (DMA)

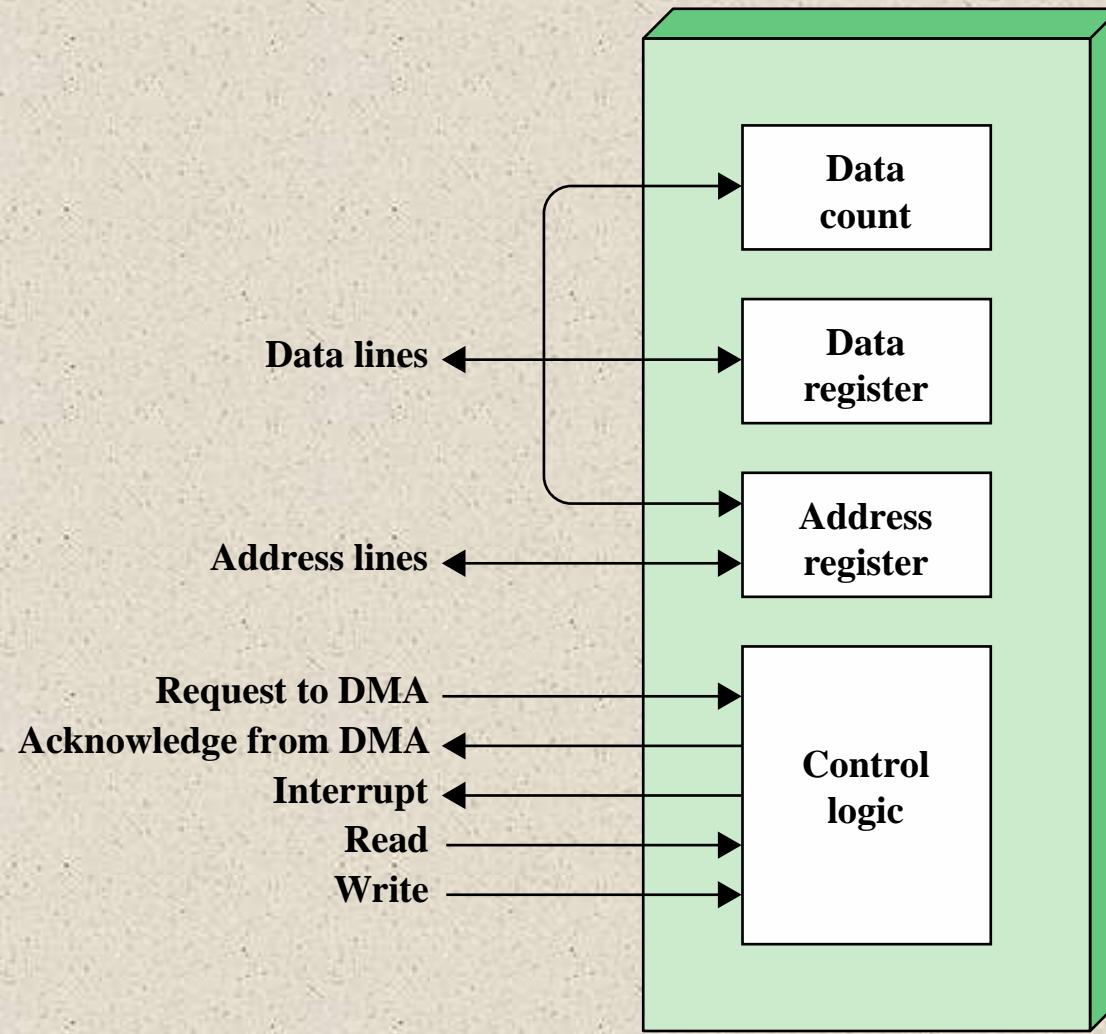


Figure 7.12 Typical DMA Block Diagram

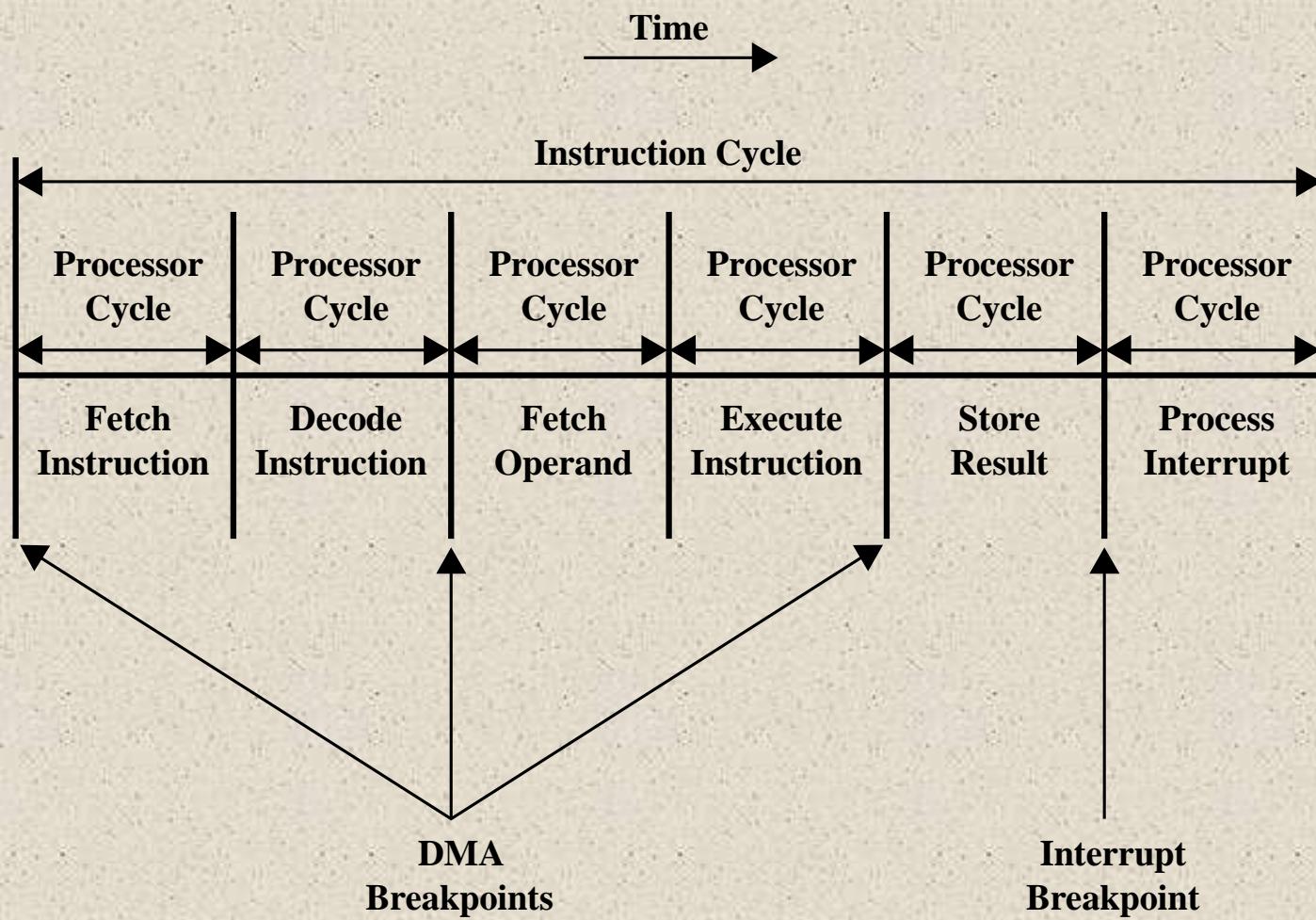


Figure 7.13 DMA and Interrupt Breakpoints During an Instruction Cycle

Controlador DMA Fly-By

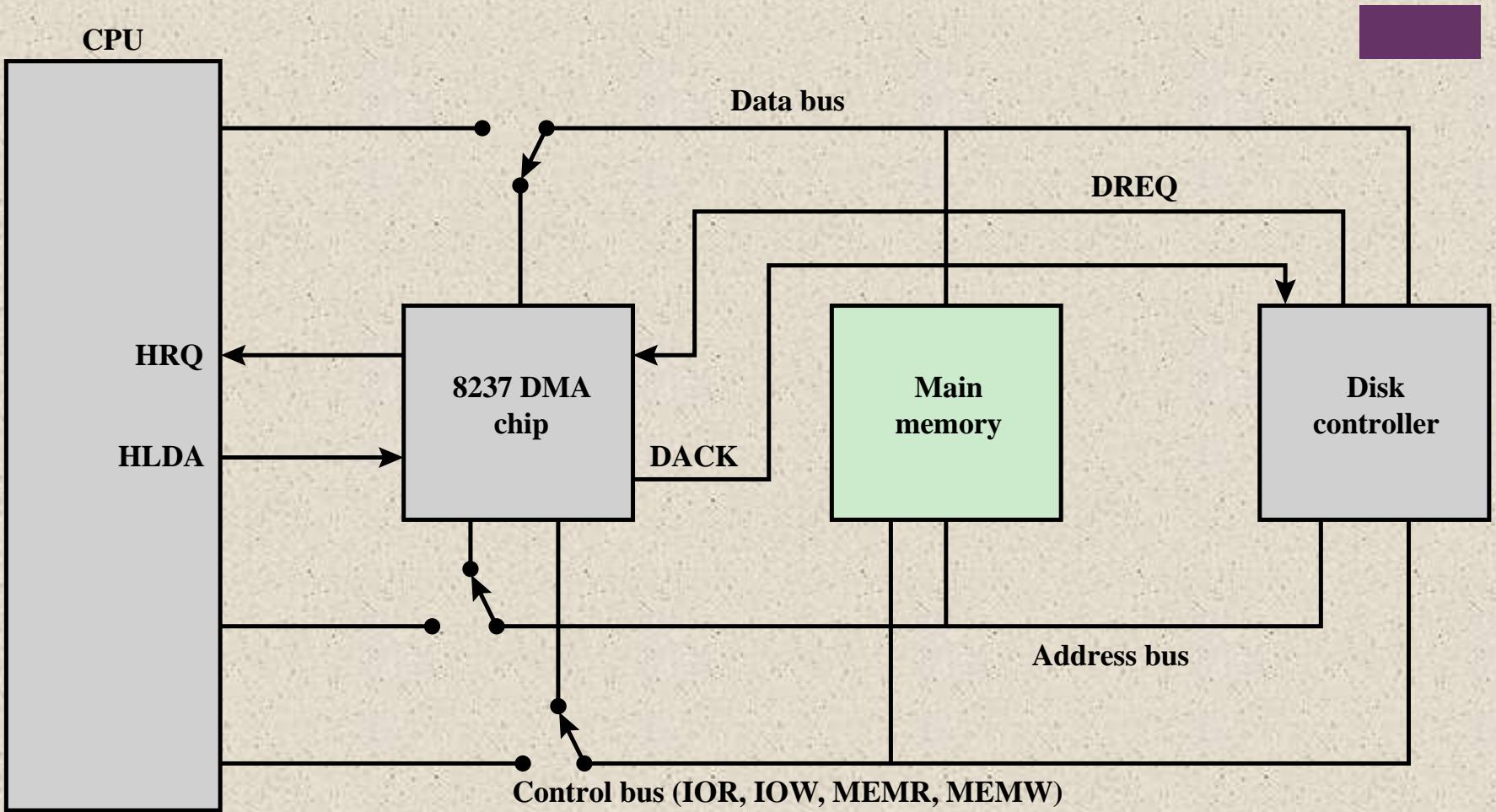
Os dados não passam
e não são
armazenados no chip
DMA

- DMA somente entre porta de E/S e local de memória
- Não entre duas portas de E/S ou dois locais de memória

Pode fazer
transferência de
memória para
memória via registro

CI 8237 contém
quatro canais DMA

- Programado de forma independente
- Qualquer um ativo
- Numerados 0, 1, 2 e 3



DACK = DMA acknowledge

DREQ = DMA request

HLDA = HOLD acknowledge

HRQ = HOLD request

Figure 7.15 8237 DMA Usage of System Bus



Acesso Direto ao Cache (Direct Cache Access - DCA)

- O DMA tradicional não consegue atender à demanda crescente devido aos aumentos nas taxas de dados para E/S de interfaces de rede
- A demanda vem principalmente da implantação de interfaces Ethernet de 10 Gbps e 100 Gbps para lidar com grandes transferências de dados de e para servidores e sistemas de alto desempenho
- Outra fonte de tráfego intenso é o Wi-Fi na faixa de gigabit
- Dispositivos de rede Wi-Fi que lidam com 3,2 Gbps e 6,76 Gbps estão se tornando amplamente disponíveis e gerando aumento de desempenho dos sistemas

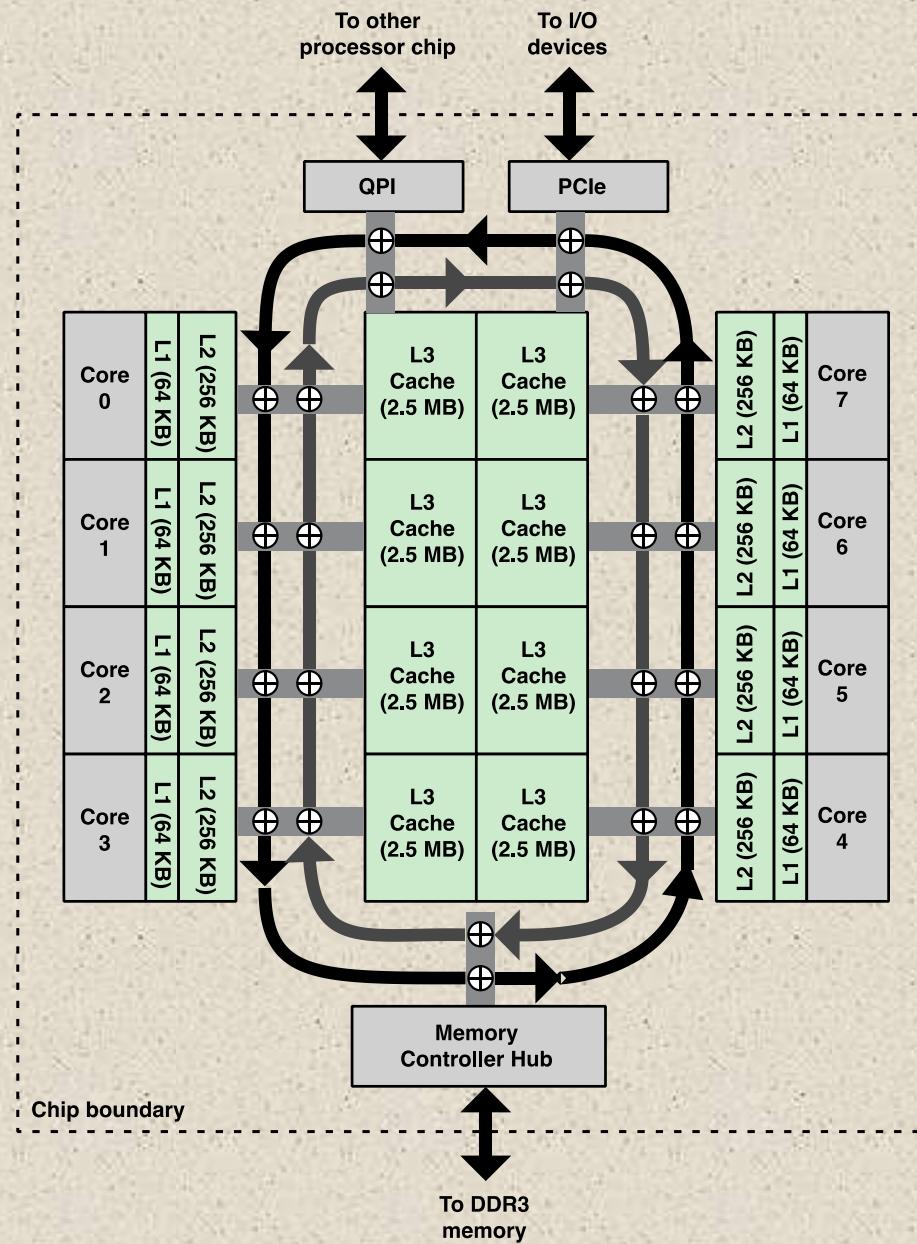


Figure 7.16 Xeon E5-2600/4600 Chip Architecture



Problemas de desempenho relacionados ao cache



Tanto para o tráfego de saída quanto para o de entrada, o núcleo, a memória principal e o cache estão todos envolvidos

Em um esquema DMA, quando um aplicativo deseja transmitir dados para interface de rede, ele coloca esses dados em um buffer atribuído pelo aplicativo na memória principal

- O núcleo transfere isso para um buffer do sistema na memória principal e cria os cabeçalhos TCP/IP necessários, que também são armazenados em buffer na memória do sistema
- O pacote é então captado via DMA para transferência via NIC
- Esta atividade envolve não apenas a memória principal, mas também o cache
- Transferências semelhantes entre os buffers do sistema e do aplicativo para o tráfego de entrada

+

Etapas do tráfego de pacotes:

Entrada

- O pacote chega
- DMA
- NIC interrompe o host
- Recupera descritores e cabeçalhos
- Ocorre uma falha de cache (miss)
- O cabeçalho é processado
- Carga útil transferida

Saída

- Transferência de pacote solicitada
- Pacote criado
- Operação de saída invocada
- Transferência DMA
- Conclusão dos sinais NIC
- O driver libera o buffer



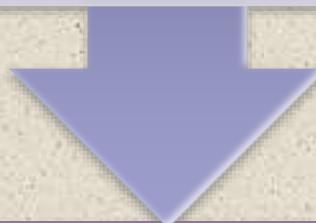
Estratégias acesso direto ao cache

A estratégia mais simples foi implementada como um protótipo em vários processadores Intel Xeon entre 2006 e 2010

Esta forma de DCA se aplica somente ao tráfego de rede de entrada

A função DCA no controlador de memória envia uma pré-busca ao núcleo assim que os dados estiverem disponíveis na memória do sistema

Isso permite que o núcleo busque rapidamente o pacote de dados do buffer do sistema



Ganhos muito mais substanciais podem ser obtidos evitando completamente o buffer do sistema na memória principal

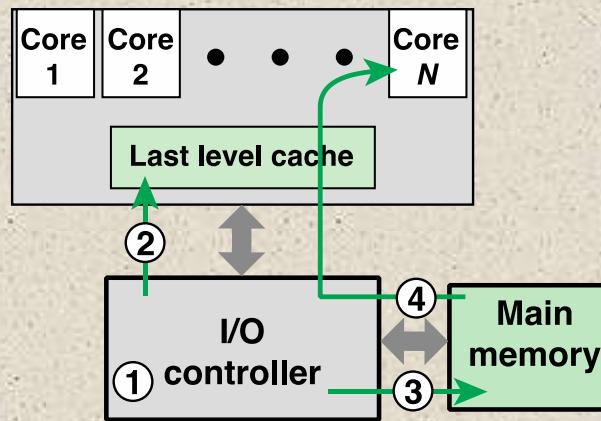
As informações do pacote e do descriptor do pacote são acessadas apenas uma vez no buffer do sistema pelo núcleo

Para pacotes de entrada, o núcleo lê os dados do buffer e transfere a carga útil do pacote para um buffer de aplicativo

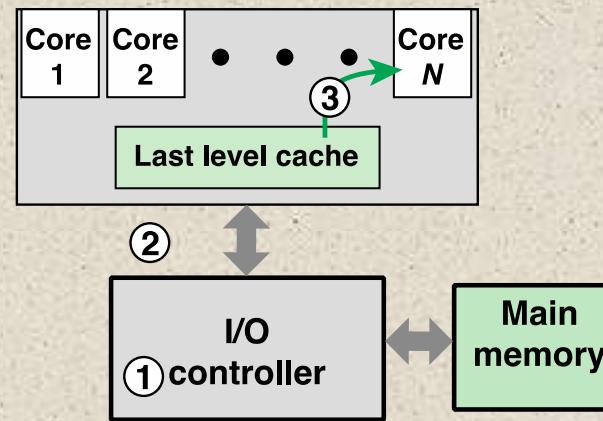
Não há necessidade de acessar esses dados no buffer do sistema novamente

Realiza a injeção de cache

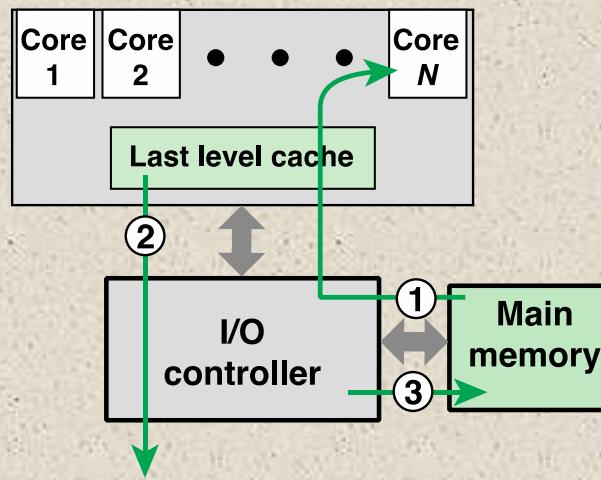
Implementado na linha de processadores Xeon da Intel, conhecido como Direct Data I/O



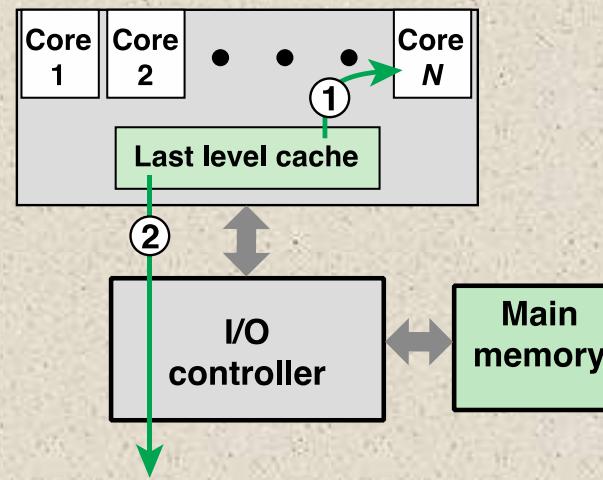
(a) Normal DMA transfer to memory



(b) DDIO transfer to cache



(c) Normal DMA transfer to I/O



(d) DDIO transfer to I/O

Figure 7.17 Comparison of DMA and DDIO

Evolução da função de E/S

1. A CPU controla diretamente um dispositivo periférico.
2. Um controlador ou módulo de E/S é adicionado. A CPU usa E/S programada sem interrupções.
3. A mesma configuração da etapa 2 é usada, mas agora com interrupções. A CPU não precisa gastar tempo esperando uma operação de E/S ser realizada, aumentando assim a eficiência.
4. O módulo de E/S recebe acesso direto à memória via DMA. Agora ele pode mover um bloco de dados para ou da memória sem envolver a CPU, exceto no início e no fim da transferência.
5. O módulo de E/S é aprimorado para se tornar um próprio processador, com um conjunto de instruções especializado e adaptado para E/S
6. O módulo de E/S tem uma memória local própria e é, de fato, um computador por si só. Com essa arquitetura, um grande conjunto de dispositivos de E/S pode ser controlado com envolvimento mínimo da CPU.

+ Barramento Serial Universal (USB)

- Amplamente utilizado para conexões periféricas
- É a interface padrão para dispositivos de velocidade mais lenta
- Interface E/S de alta velocidade comumente usada
- Passou por várias gerações
 - USB 1.0
 - Definiu uma taxa de dados de *baixa velocidade* de 1,5 Mbps e uma taxa de *velocidade máxima* de 12 Mbps
 - USB 2.0
 - Fornece uma taxa de dados de 480 Mbps
 - USB 3.0
 - Barramento de alta velocidade chamado *SuperSpeed* em paralelo com o barramento USB 2.0
 - Velocidade de sinalização do *SuperSpeed* é de 5 Gbps , mas devido à sobrecarga de sinalização a taxa de dados utilizável é de até 4 Gbps
 - USB 3.1
 - Inclui um modo de transferência mais rápido chamado *SuperSpeed +*
 - Este modo de transferência atinge uma taxa de sinalização de 10 Gbps e uma taxa de dados teórica utilizável de 9,7 Gbps
- É controlado por um controlador host raiz que se conecta aos dispositivos para criar uma rede local com uma topologia de árvore hierárquica

+

SCSI (Small Computer System Interface)

- Um padrão outrora comum para conectar dispositivos periféricos a computadores de pequeno e médio porte
- Perdeu popularidade para USB e FireWire nos sistemas menores
- Versões de alta velocidade continuam populares para suporte de memória em massa em sistemas empresariais
- A organização física é um barramento compartilhado, que pode suportar até 16 ou 32 dispositivos, dependendo da geração do padrão
 - O barramento permite transmissão paralela em vez de serial, com uma largura de barramento de 16 bits nas gerações anteriores e 32 bits nas gerações posteriores.
 - As velocidades variam de 5 Mbps na especificação SCSI-1 original a 160 Mbps no SCSI-3 U3





InfiniBand

- Especificação de E/S voltada para o mercado de servidores
- A primeira versão foi lançada no início de 2001
- Altamente confiável e implementada na série de mainframes IBM zEnterprise
- O padrão descreve uma arquitetura e especificações para fluxo de dados entre processadores e dispositivos de E/S inteligentes
- Tornou-se uma interface popular para redes de área de armazenamento e HPC
- Permite que servidores, armazenamento remoto e outros dispositivos de rede sejam conectados em uma estrutura central de switches e links
- A arquitetura baseada em switch pode conectar até 64.000 servidores, sistemas de armazenamento e dispositivos de rede

PCI Express

- Sistema de barramento de alta velocidade para conectar periféricos de uma ampla variedade de tipos e velocidades



SATA

- Serial Advanced Technology Attachment
- Uma interface para sistemas de armazenamento em disco
- Fornece taxas de dados total de até 6 Gbps , com um máximo de 300 Mbps por dispositivo
- Amplamente utilizado em computadores de mesa e em aplicações industriais e embarcadas



Ethernet



- Tecnologia de rede com fio predominante
- Evoluiu para suportar taxas de dados de até 100 Gbps e distâncias de alguns metros a dezenas de km
- Tornou-se essencial para dar suporte a computadores pessoais, estações de trabalho, servidores e dispositivos de armazenamento de dados em massa em organizações grandes e pequenas
- Começou como um sistema experimental de 3 Mbps baseado em barramento com cabo coaxial
- Passou de baseado em barramento para baseado em switch
 - A taxa de dados aumentou regularmente em uma ordem de magnitude
 - Existe um comutador central com todos os dispositivos conectados diretamente em estrela
- Os sistemas Ethernet estão atualmente disponíveis em velocidades de até 100 Gbps

+

Wi-fi

- A tecnologia predominante de acesso à Internet sem fio
- Permite conectar computadores, tablets, smartphones e outros dispositivos eletrônicos, como câmeras de vídeo, TVs e termostatos
- Na empresa tornou-se um meio essencial para aumentar a produtividade dos trabalhadores e a eficácia da rede
- Os hotspots públicos expandiram-se drasticamente para fornecer acesso à Internet na maioria dos locais públicos

- À medida que a tecnologia de antenas, técnicas de transmissão sem fio e design de protocolo sem fio evoluíram, o comitê IEEE 802.11 conseguiu introduzir padrões para novas versões de Wi-Fi em velocidades cada vez mais altas.
- A versão atual é 802.11ax ou WiFi 6e (2021) com uma taxa de dados de 9,6 Gbps



+

Resumo

Capítulo 7

- Dispositivos externos
 - Teclado/monitor
 - Unidade de disco
- Módulos de E/S
 - Função do módulo
 - Estrutura do módulo de E/S
- E/S programada
 - Visão geral de E/S programadas
 - Comandos/instruções de E/S
- Acesso direto à memória
 - Desvantagens da E/S programada e controlada por interrupção
 - Função DMA
 - Controlador Intel 8237A DMA

Entrada/Saída

- E/S acionada por interrupção
 - Interromper processamento
 - Problemas de design
 - Controlador de interrupção Intel 82C59A
 - Interface periférica programável Intel 82C55A
- Acesso direto ao cache
 - DMA usando cache de último nível
 - Problemas de desempenho relacionados ao cache
 - Estratégias de acesso direto ao cache
 - E/S de dados direta
- Canais de E/S e processadores
 - A evolução da função I/O
 - Características dos canais de E/S