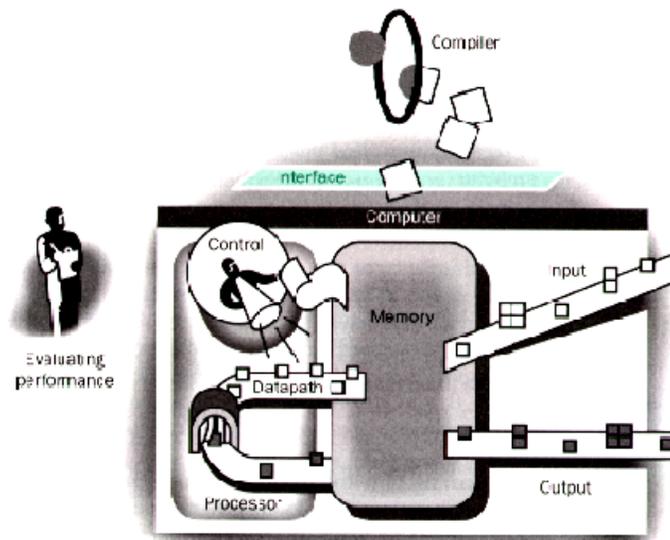


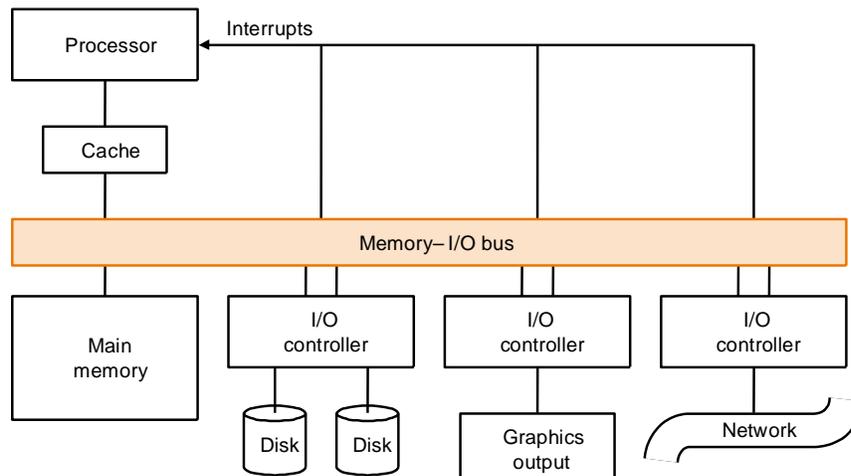
Interligando Processador e Periféricos



Sistema de E/S

- *Dependentes da tecnologia*
- *Considerações de projeto:*
 - *a possibilidade de expandir o sistema*
 - *o comportamento no caso de falhas*
 - *o desempenho*

Sistema de E/S



Sistema de E/S

■ *Desempenho*

- *tempo de acesso (latência)*
- *taxa de transferência (throughput)*
- *Dependência:*
 - *características do dispositivo*
 - *conexão com o sistema*
 - *hierarquia de memória*
 - *sistema operacional (software de I/O)*

Desempenho com E/S

- **Aplicação:**
 - $tempo_{execução} = 100 \text{ seg}$
 - CPU = 90 seg, E/S = 10 seg
 - $tempo_{execução} = tempo_{CPU} + tempo_{E/S}$
- **desempenho da CPU melhora de 50% a cada ano e o do sistema de E/S não**
 - $tn = to / (1+c)^n$
- **Quanto mais rápida será a aplicação daqui a 5 anos?**

Desempenho com E/S

N anos	Tempo _{CPU}	Tempo _{E/S}	Tempo _{total}	% tempo _{I/O}
0	90	10	100	10
1	90/1.5=60	10	70	14
2	40	10	50	20
3	27	10	37	27
4	18	10	28	36
5	12	10	22	45

Melhora do Desempenho_{CPU}=7.5

Melhora do Tempo_{Total}=4.5

Desempenho de E/S

■ *Métricas*

- *depende da aplicação*
 - *Taxa de transferência*
 - *Quantidade de dado transferido*
 - *Quantidade de operações de E/S*
 - *Tempo de resposta*

Dispositivos de E/S

■ *Comportamento*

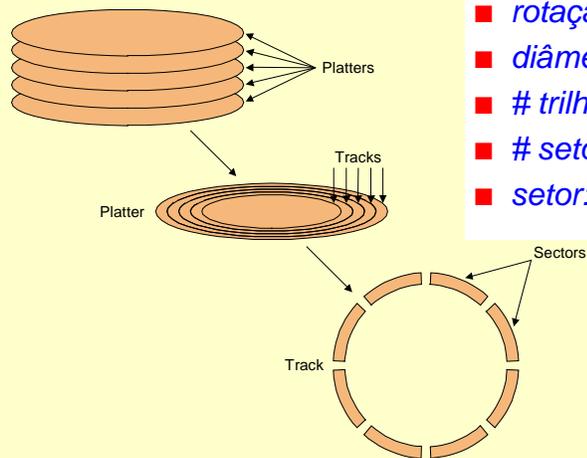
- *Entrada (read once)*
- *Saída (write only)*
- *Armazenamento(Podem ser re-lidos e re-escritos)*

■ *Taxa de Transferência de Dados*

Dispositivos de E/S

Dispositivo	Utilização	Interface	Taxa (KB/s)
Teclado	Entrada	Humana	0.01
Mouse	Entrada	Humana	0.02
Scanner	Entrada	Humana	200
Impressora. de linha	Saída	Humana	1
Impressora laser	Saída	Humana	100
Network	Entrada/Saída	Máquina	10000
Disco Magnético	Armazenamento	Máquina	2000
Disco Ótico	Armazenamento	Máquina	500

Discos magnéticos



- # discos: 1-15 (2 lados)
- rotação: 3600 - 7200 rpm
- diâmetro: 1-8 inch
- # trilhas: 1000-5000
- # setores: 64-200
- setor: 512 bytes

Discos magnéticos

■ Acesso:

- *posicionamento da cabeça leitura/escrita*
 - *seek time (mínimo, máximo, médio)*
- *seleção do setor*
 - *rotational delay (médio)*
 - $ARD = 0.5/3600 \text{ RPM} = 0.5/3600/60 = 8.3\text{ms}$
- *tempo de transferência*
 - *depende do tamanho do setor, rotação e capacidade de armazenamento da trilha*
 - *1997: 2 - 15 MB/seg*
 - *cache de disco: 40 MB/seg*

Discos magnéticos

■ Calculando tempo de leitura:

- *Qual o tempo para ler um setor de um disco rotacionando a 5400 rpm?*
 - *Seek time = 12ms (médio)*
 - *Taxa = 5 MB/seg*
 - *Overhead controle = 2ms*

Tempo = seek + rotational + transferencia + overhead

Tempo = 12ms + 0.5/(5400/60) + 0.5KB/(5MB/seg) + 2ms

Tempo = 12ms + 5.6 + 0.1 + 2ms = 19.7ms

Discos magnéticos

■ E se o seek_time for 25% do descrito?

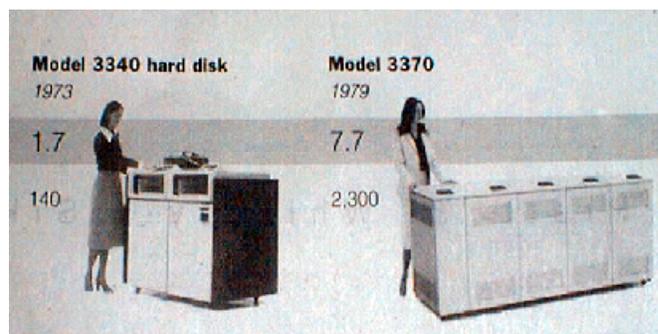
- Rotação= 5400 RPM
- Seek time = 3ms (médio)
- Taxa = 5 MB/seg
- Overhead controle = 2ms

Tempo = seek + rotational + transferencia + overhead

Tempo = 3ms + 0.5/(5400/60) + 0.5KB/(5MB/seg) + 2ms

Tempo = 3ms + 5.6 + 0.1 + 2ms = 10.7ms

Discos / História

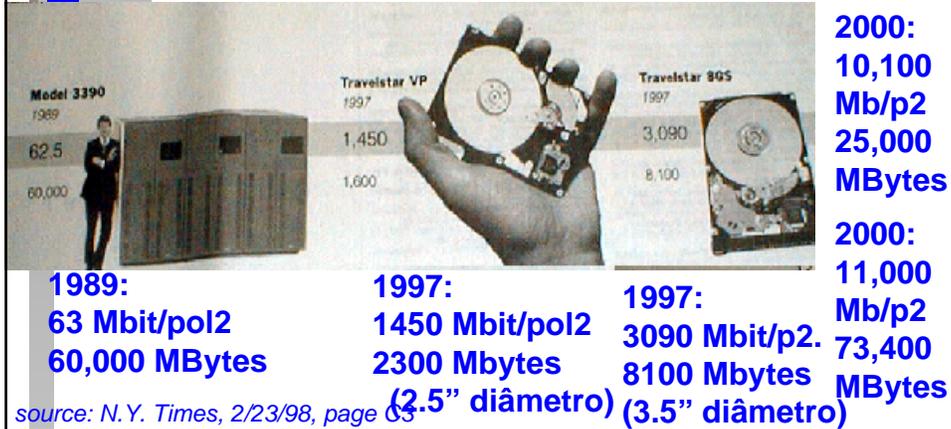
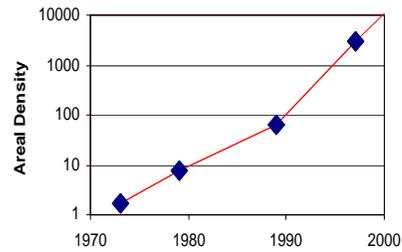


1973:
1.7 Mbit/pol2
140 MBytes

1979:
7.7 Mbit/pol2
2,300 MBytes

source: New York Times, 2/23/98, page C3,
"Makers of disk drives crowd even more data into even smaller spaces"

Discos: história



Estado da Arte: Ultrastar 72ZX



- 73.4 GB, 3.5 inch disk
- 2 ϕ /MB
- 16 MB track buffer
- 11 platters, 22 surfaces
- 15,110 cylinders
- 7 Gbit/sq. in. areal density
- 17 watts (idle)
- 0.1 ms controller time
- 5.3 ms avg. seek
(seek 1 track => 0.6 ms)
- 3 ms = 1/2 rotation
- 37 to 22 MB/s to media

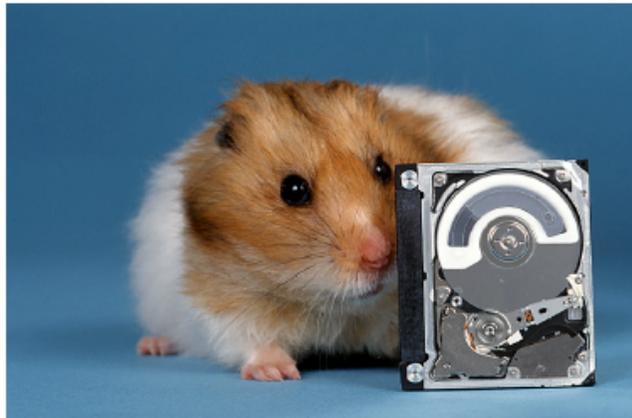
source: www.ibm.com;
www.pricewatch.com; 2/14/00

Discos: Limites

- Capacidade crescente (60%/ano) and Taxa de tranf. crescente (40%/ano)
- Redução lenta do tempo de acesso (8%/ano)
- Tempo para leitura completa:

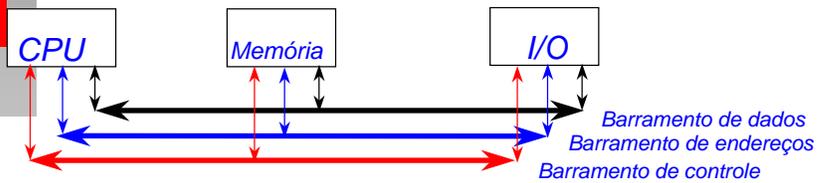
Ano	Sequencial	Randômica
1990	4 minutos	6 horas
2000	12 minutos	1 semana

Estado da Arte



1 GBytes

Barramentos



Vantagens:

- versatilidade
- baixo custo

Desvantagens:

- limita taxa de transmissão
- # de dispositivos
- tamanho do barramento

Barramentos

■ Como aumentar o desempenho?

- diminuir tempo de resposta
 - minimizar tempo de acesso
- aumentar taxa de transferência
 - usar buffers
 - transferência de grandes blocos

Comunicação via barramento

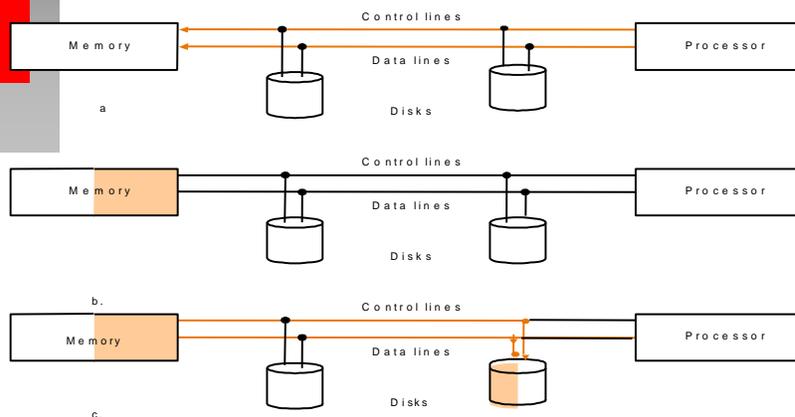
■ Linhas:

- controle
- dado (endereços e dados)

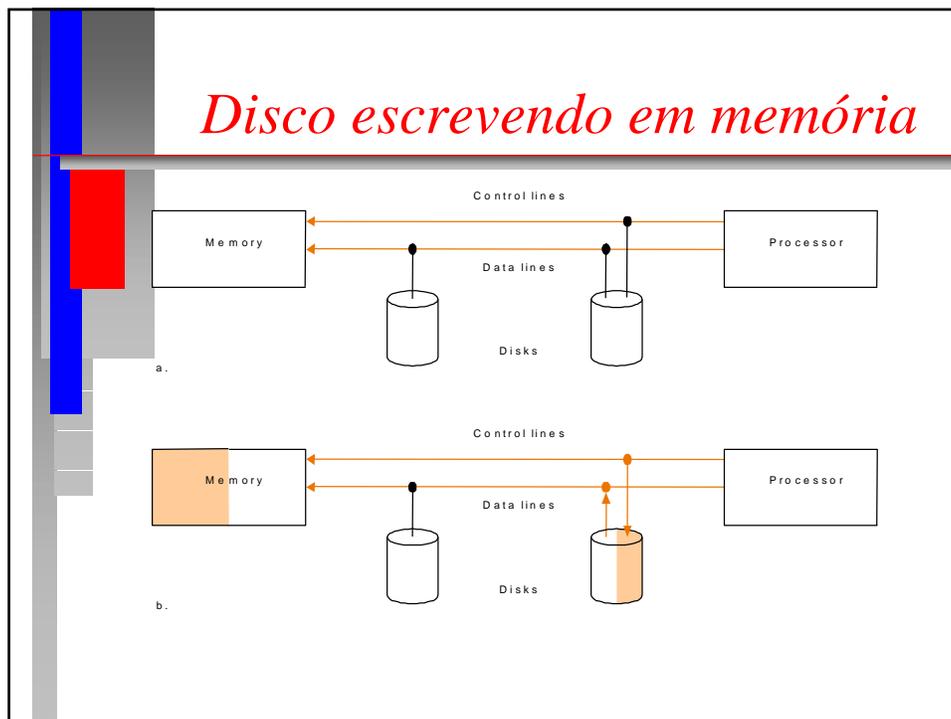
■ Transação

- envio endereço
- envio/recebimento de dado
- leitura (output)
 - memória → CPU-I/O
- escrita (input)
 - memória ← CPU-I/O

Disco lendo memória



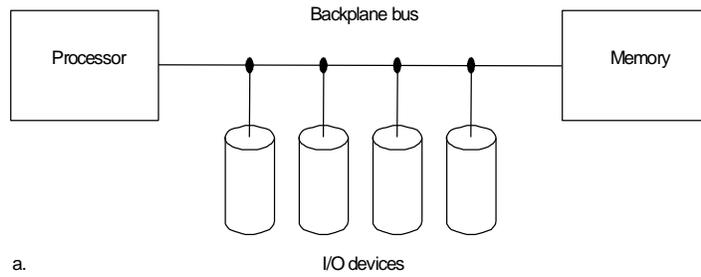
Disco escrevendo em memória



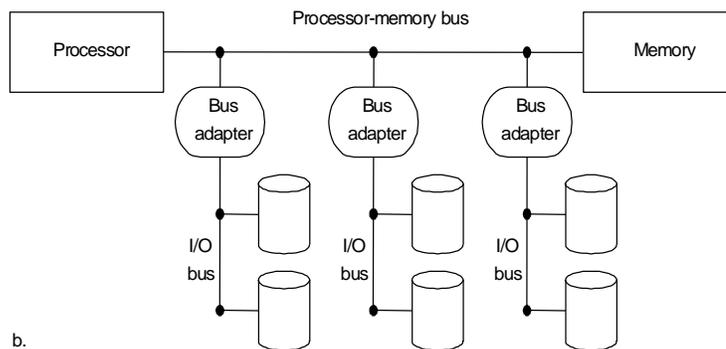
Tipos de barramentos

- **Processador-memória**
 - curtos e de alta velocidade
 - específicos
- **I/O**
 - longos
 - compatíveis com vários dispositivos (padronizados)
- **Backplane (CPU, memória, I/O)**
 - padronizados
 - baixo custo

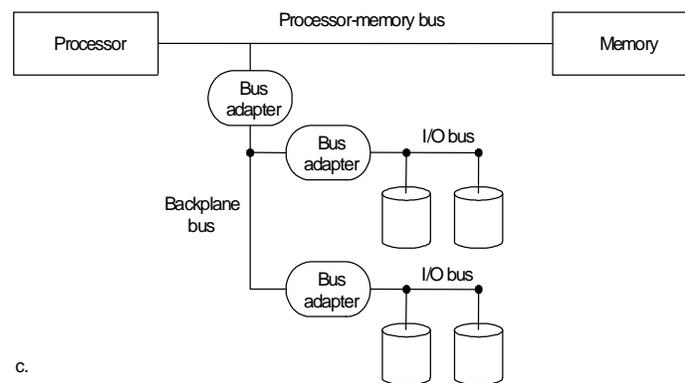
Tipos de barramentos



Tipos de barramentos



Tipos de barramentos



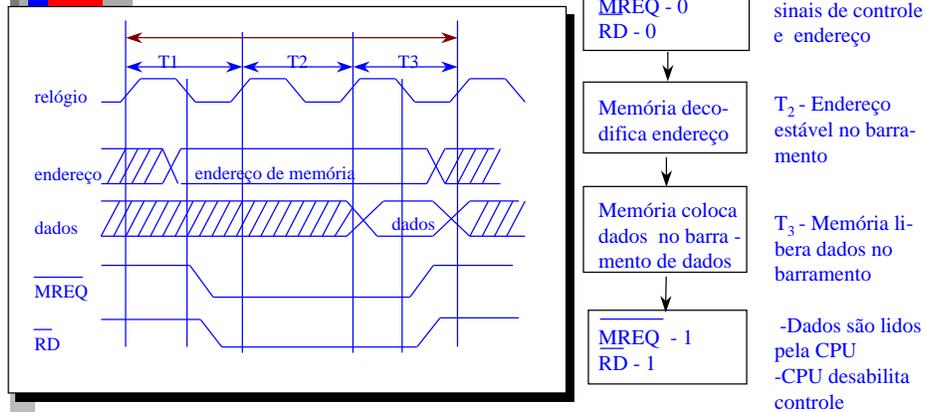
Barramento - Temporização

■ Barramento síncrono

- Este tipo de barramento exige que todo o tráfego de dados e controle seja sincronizado sob uma mesma base de tempo (clock)
- Vantagens:
 - simplicidade
 - desempenho
- Problemas:
 - dispositivos devem ter a mesma frequência do clock
 - curtas distancias
- Processador-memória

Barramento - Síncrono

✓ Ciclo de leitura

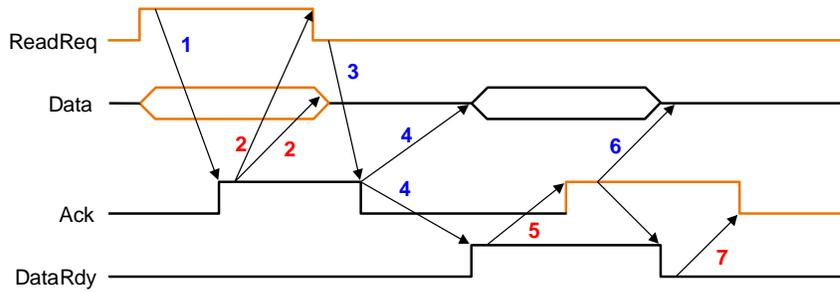


Barramento - Temporização

■ Barramento assíncrono

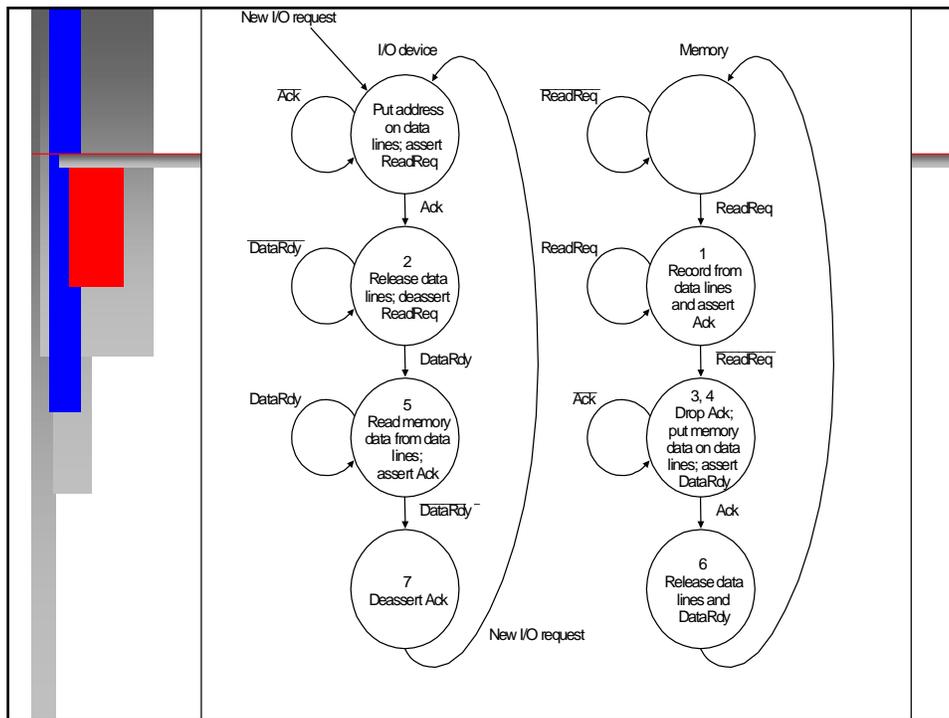
- Este barramento não possui um relógio mestre. Os ciclos de barramentos podem ter qualquer duração e não precisam ser iguais para todas as situações.
- Podem incluir vários dispositivos
- Necessita de protocolo de handshake
- Pode ser longo
- Barramento de I/O

Barramento assíncrono



Dispositivo de E/S

Memória



Síncrono vs. Assíncrono

■ Síncrono:

- tempo de ciclo = 50 ns
- transmissão => 1 ciclo

■ Assíncrono

- handshake => 40ns

■ Dados: 32 bits

■ Memória : 200 ns (leitura)

■ Qual a taxa para leitura de 1 palavra?

Síncrono vs. Assíncrono

■ Síncrono:

- envia endereço => 50ns
- lê memória => 200 ns
- envia dado para dispositivo => 50ns
- taxa_{sinc.} = 4bytes/300 ns = 13.3 MB/seg

■ Assíncrono

- 1 => 40ns
- 2,3,4 => max. (3x40, 200) = 200ns
- 5,6,7 => 3x40= 120 ns
- taxa_{assinc.} = 4 bytes/360ns = 11.1 MB/seg

Melhorando o desempenho

- *Aumentar barramento de dados*
- *Separar linhas de dados e endereços*
- *Transferência em blocos*



Maior custo
Maior tempo de resposta

Melhorando o desempenho

- *Barramento síncrono:*
 - *dado => 64 bits*
 - *clock = 200 MHz*
 - *transferência => 1 ciclo*
 - *idle time => 2 ciclos*
- *Memória:*
 - *primeiras 4 palavras => 200ns*
 - *4 palavras adicionais => 20ns*
 - *sobreposição: transf. de dado lido e leitura de próximo bloco*

Melhorando o desempenho

- Qual a latência e a taxa de transferência para se ler 256 palavras usando:
 - blocos de 4 palavras de 32 bits
 - blocos de 16 palavras de 32 bits

?

Melhorando o desempenho

- Blocos de 4 palavras
 - 1 ciclo para enviar endereço
 - $200\text{ns} / 5\text{ns} = 40$ ciclos para leitura
 - 2 ciclos para envio de dado
 - 2 idle cycles



45 ciclos/bloco

$\#ciclos_{total} = 256/4 \times 45 = 2880$ ciclos

Latencia = $2880 \times 5 \text{ ns} = 14.400 \text{ ns}$

$\#transações/seg = 64 \times (1/14400) = 4,44 \text{ M transações/seg}$

Taxa = $(256 \times 4) \text{ bytes} \times 1/14400 \text{ seg} = 71,11 \text{ MB/seg}$

Melhorando o desempenho

■ Blocos de 16 palavras

- 1 ciclo para enviar endereço
- $200\text{ns} / 5\text{ns} = 40$ ciclos para leitura das 4 pal.
- 2 ciclos para envio de dado e leitura das prox. 4 pal.
- 2 idle cycles onde a leitura é terminada



$$1+40+(2+2) + (2+2) + (2+2) + (2+2)= 57 \text{ ciclos/bloco}$$

$$\#ciclos_{total} = 256/16 \times 57 = 912 \text{ ciclos}$$

$$\text{Latencia} = 912 \times 5 \text{ ns} = 4560 \text{ ns}$$

$$\#transações/seg = 16 \times (1/4560) = 3.51\text{M transações/seg}$$

$$\text{Taxa} = (256 \times 4) \text{ bytes} \times 1/4560 \text{ seg} = 224,56 \text{ MB/seg}$$

Acessando o barramento

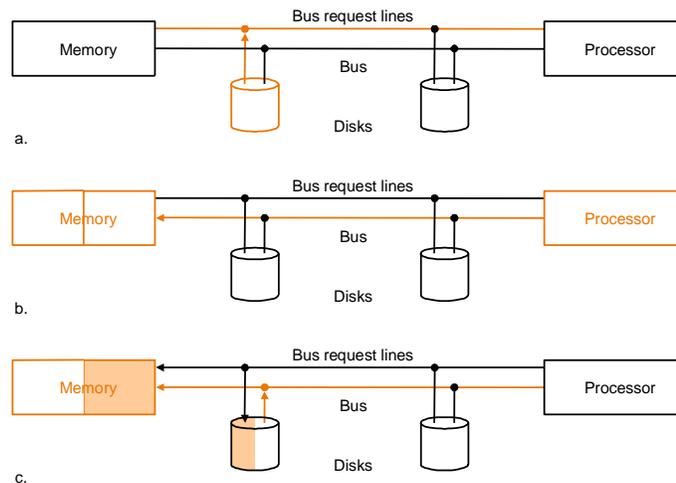
■ Mestre do barramento

- inicia e controla todas as requisições ao barramento

■ Quem pode ser mestre?

- Processador
- Dispositivo de I/O

Mestre único (processador)



Mestre único (processador)

■ *Vantagem:*

- *simplicidade*

■ *Desvantagem:*

- *Processador coordena todas as transações de barramento*
- *degradação do desempenho*

Múltiplos Mestres

- *Processador e dispositivos de E/S podem se tornar mestres*
- *Necessidade de protocolo*
 - *request*
 - *granted*
 - *bus-release*
- *Necessidade de esquema de arbitragem*

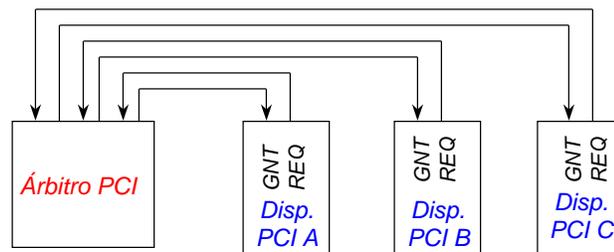
Barramento - Arbitragem

- *O que acontece quando dois ou mais dispositivos querem se tornar mestres do barramento ao mesmo tempo?*
 - *A arbitragem decide qual mestre terá o controle do barramento num dado instante*
 - *Arbitragem centralizada*
 - *Arbitragem descentralizada*

Barramento - Arbitragem Centralizada

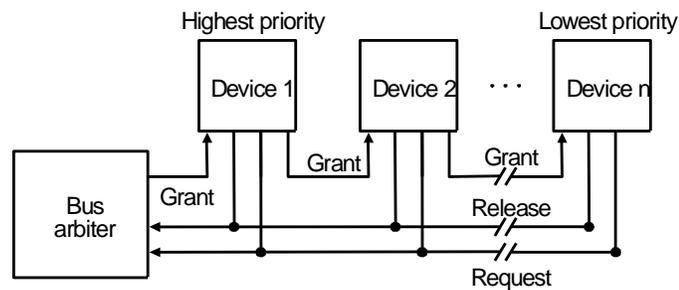
☑ Arbitragem no barramento PCI (centralizado)

- Exemplo para três dispositivos
- O árbitro decide qual mestre controla o barramento
- Desvantagem:
 - dependência do árbitro



Barramento - Arbitragem Híbrida (centralizada e distribuída)

■ Barramento de um nível usando daisy-chaining

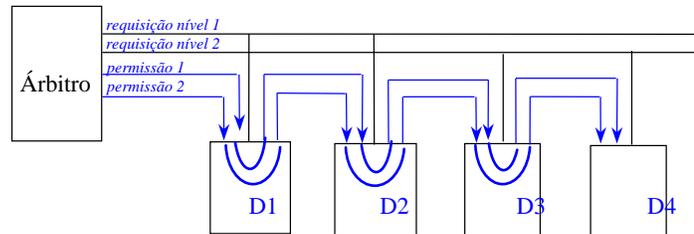


Características

1. Todos os dispositivos são ligados em série, assim a permissão, dada pelo árbitro, pode ou não se propagar através da cadeia.
2. Cada dispositivo deve solicitar acesso ao barramento.
3. O dispositivo mais próximo do árbitro tem maior prioridade.

Barramento - Arbitragem Híbrida (centralizada e distribuída)

☑ Arbitragem com dois níveis de prioridade



Características

1. Cada dispositivo se liga a um dos níveis de requisição.
2. Os dispositivos com tempos mais críticos se ligam aos níveis de maior prioridade.
3. Se múltiplos níveis de prioridade são requeridos ao mesmo tempo, o árbitro solta a permissão apenas para os de prioridade mais alta.

Barramento - Arbitragem Descentralizada

- **Arbitragem distribuída por seleção própria**
 - múltiplas linhas de requisição
 - dispositivos requisitantes determinam quem vai usar o barramento (prioridade)
 - Ex: NuBus (Macintosh)
- **Arbitragem distribuída com detecção de colisão**
 - requisições independentes
 - colisões são detectadas e resolvidas
 - Ex: Ethernet

Barramento - Aspectos de projeto

■ Considerações na implementação de do barramento do sistema:

Opções	Alta Performance	Baixo custo
Largura do barramento	Endereços e dados separados	Multiplexação das linhas de endereço/dados
Largura de dados	16 bits, 32 bits, ... (Quanto maior mais rápido)	Menor, mais barato
Transações por pacotes	Múltiplas palavras menos overhead	Transferência por palavra é mais simples
Barramentos masters	Múltiplos mestres (requer arbitragem)	Apenas um mestre (sem arbitragem)
Relógio (Clocking)	Síncrono	Assíncrono

Barramento - Padrões

Característica	PCI	SCSI
Tipo	Backplane	I/O
Largura (dados)	32-64	8-32
Multiplexação end/dados?	multiplexado	Multiplexado
Número mestres	múltiplos	Múltiplos
Arbitragem	Centralizada	Self-selection
Relógio	Síncrono(33-66 MHz)	Assíncrono ou síncrono(5-10MHz)
Bandwith(teórica)	133-512 MB/seg	5-40MB/Seg
Bandwith(medida)	80 MB/seg	2.5 – 40 ou 1.5 MB/Seg(ass.)
Número de dispo.	1024	7-31
Comprimento	0.5 metro	25 metros
Nome	PCI	ANSI X3.31

Executando operações de E/S

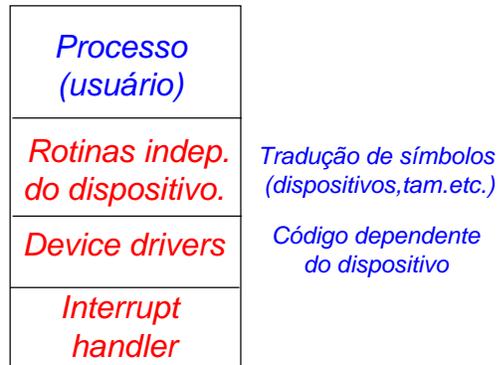
- *Comunicação com os dispositivos de E/S:*
 - *envio de comandos aos dispositivos*
 - *transferência de dados de/para dispositivos*
 - *análise do status dos dispositivos e da transmissão*
- *Interface entre usuário e dispositivo*
 - *Sistema Operacional*

Suporte do S.O. a E/S

- *Proteção*
 - *garante o acesso a dispositivos/dados para os quais se tenha permissão*
- *Abstração (dispositivo)*
 - *possui rotinas específicas com detalhes de cada dispositivo*
- *Gerenciamento*
 - *trata as interrupções causadas pelos dispositivos*
- *Escalonamento*
 - *controla a utilização de dispositivos compartilhados entre processos*

Software de E/S

■ *Organizado em camadas*



Implementando E/S

■ *Instruções especiais*

- *Registradores especiais*
 - *status dos dispositivos*
 - *dado a ser recebido/transmitido*
- *Proteção*
 - *instruções privilegiadas (só podem ser executadas via S.O.)*

Implementando E/S

■ *E/S mapeada em memória*

- *Área de memória para armazenar status e dados*
- *Instruções de E/S : instruções de acesso à memória*
- *Proteção:*
 - *tradução do endereço*

Implementando E/S

■ *Pooling*

CPU

- 1- le status do disp.*
- 3- Inspecciona status, se não está pronto va para 1*
- 4- escreve no reg. dado*
- 6- se existe mais dados va para 1*

DISPOSITIVO

- 2- envia status*
- 5- aceita dado, status=ocupado até escrita terminada*

Pooling

■ *Vantagens:*

- *E/S controlada pela CPU*
- *simplicidade de implementação*
- *Dispositivos que iniciam E/S:*
 - *Mouse*
 - *Impressora*

■ *Problema:*

- *CPU fica esperando pelo dispositivo*

Pooling - Overhead

■ $\#ciclos_{pooling}=400$, *clock= 500 MHz*

■ *Qual a fração de tempo da CPU em cada um dos casos:*

- *Mouse deve ser verificado 30 vezes/seg*
- *Floppy transfere palavras de 16 bits a uma taxa de 50 KB/seg*
- *O disco rígido transfere blocos de 4 palavras a uma taxa de 4MB/seg.*

?

Pooling - Overhead

■ Mouse:

- $\#ciclos = 30 \times \#ciclos_{pooling} = 30 \times 400 = 12000 \text{ ciclos}$
- $Fra\c{c}o_{CPU} = 12000 \times 10^2 / 500 \times 10^6 = 0.002\%$

■ Floppy:

- $(50 \text{ KB/seg}) / 2 \text{ bytes} = 25\text{K pooling/seg}$
- $\#ciclos = 25\text{K} \times 400 = 10 \times 10^6$
- $Fra\c{c}o_{CPU} = 10 \times 10^6 \times 10^2 / 500 \times 10^6 = 2\%$

Pooling - Overhead

■ Disco r gado:

- $(4\text{MB/seg}) / 16 \text{ bytes} = 250\text{K pooling/seg}$
- $\#ciclos = 250\text{K} \times 400 = 100 \times 10^6$
- $Fra\c{c}o_{CPU} = 100 \times 10^6 \times 10^2 / 500 \times 10^6 = 20\%$

Interrupt-driven I/O

- CPU executa outras instruções enquanto espera E/S
- Sincronização: interrupção
 - disp. E/S está pronto para nova transferência
 - operação de E/S terminou
 - ocorreu erro
- Vantagem:
 - melhor utilização da CPU

Interrupt-driven Overhead

- Exemplo anterior: Disco rígido
 - transferência em 5% do tempo
 - $\text{Overhead}_{\text{interrupção}} = 500 \text{ ciclos}$
 - $(4\text{MB}/\text{seg})/16 \text{ bytes} = 250\text{K transf./seg}$
 - $\#\text{ciclos} = 250\text{K} \times 500 = 125 \times 10^6$
 - $\text{Fração}_{\text{CPU}} = 125 \times 10^6 \times 10^2 / 500 \times 10^6 = 25\%$
como o disco só transmite em 5% do tempo:
 - $\text{Fração}_{\text{CPU}} = 25\% \times 5\% = 1.25\%$

Vantagem: Ausência de overhead quando dispositivo não está transmitindo

DMA(transf. autonoma)

- *Transferência de blocos de dados*
 - 80 a 512 bytes
- *E/S controlada pelo dispositivo de E/S*
 - *parametros:*
 - end. memória
 - tamanho blocos
 - end. Dispositivo
- *Interrupção da CPU quando DMA termina*

DMA Overhead

- *Exemplo anterior: Disco rígido*
 - Blocos => 8KB
 - transmissão em 100% do tempo
 - Setup=1000 ciclos
 - $\text{Overhead}_{\text{interrupção}} = 500 \text{ ciclos}$
 - $\text{Tempo}_{\text{DMA}} = 8\text{KB}/(4\text{MB}/\text{seg}) = 2 \times 10^{-3} \text{ seg}$
 - $\#\text{ciclos}/\text{seg} = (1000+500)/ 2 \times 10^{-3} = 750 \times 10^3$
 - $\text{Fração}_{\text{CPU}} = 750 \times 10^3 \times 10^2 / 500 \times 10^6 = 0,2\%$

Projetando um Sistema de E/S

■ *Restrições de projeto:*

- *latência*
- *taxa de transferência*

■ *Latência:*

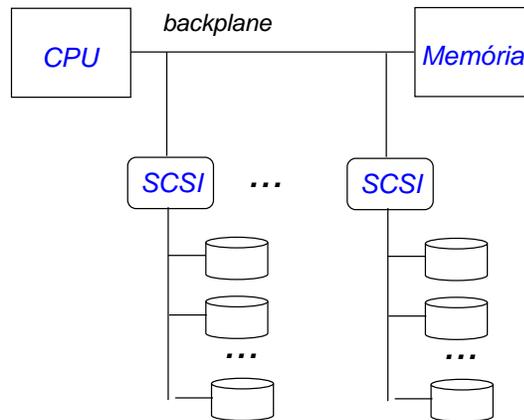
- *tempo de resposta admitido pela aplicação*
- *simplicidade:*
 - *não considera a carga do sistema*

Projetando um Sistema de E/S

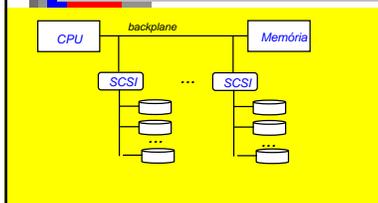
■ *Taxa de Transferência*

- *taxa do componente mais “lento”*
- *outros componentes devem ser configurados de forma a trabalhar nesta taxa*

Projetando um Sistema de E/S



Projetando um Sistema de E/S



■ CPU

- 300×10^6 instr./seg
- I/O $\Rightarrow 50 \times 10^3$ instr. do S.O

■ Bus - taxa=100 MB/seg

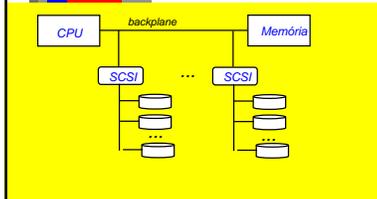
■ SCSI

- taxa=20MB/seg
- até 7 discos

■ Discos

- taxa=5MB/seg
- seek+rotational=10ms

Projetando um Sistema de E/S



■ Aplicação

- leitura 64KB
- prog. Usuário => 100×10^3 instr

Qual a taxa de I/O do sistema e quantos discos pode ter de forma que esta taxa seja mantida?

Projetando um Sistema de E/S

■ $Taxa_I/O_{CPU} = (\# \text{ instr./seg}) / (\text{instr. I/O/seg})$
 $= 300 \times 10^6 / (50 + 100) \times 10^3 = 2000 \text{ I/O/seg}$

■ $Taxa_I/O_{bus} = \text{taxa/bytes por I/O} =$
 $100 \times 10^6 / 64 \times 10^3 = 1562 \text{ I/O/seg}$



Barramento tem a menor taxa

Configurar o resto do sistema de forma a se ter esta taxa de I/O.

Projetando um Sistema de E/S

Quantos discos podemos ter no sistema?

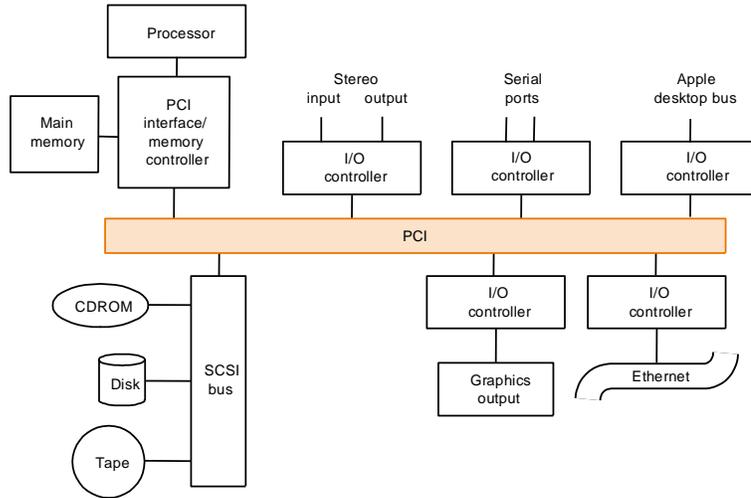
- $Tempo_I/O_{disco} = seek + rotational + transf. = 100 + 64K/5MB = 22,8 \text{ ms}$
- $Taxa_I/O_{disco} = 1/22,8ms = 43,9 \text{ I/O/seg}$
- $\# \text{ discos} = 1562/43,9 = 36 \text{ discos}$

Projetando um Sistema de E/S

Quantos barramentos SCSI?

- $\# \text{ discos} = 36 \text{ discos}$
- $Taxa_{disco} = tam.bloco/tempo = 64KB/22.8ms = 2.74MB/seg$
- $Taxa_{SCSI} = 20MB/seg \Rightarrow \text{pode ter 7 discos por barramento}$
- $\#Bus_SCSI = 36/7 = 6$

E/S Macintosh 7200



This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.