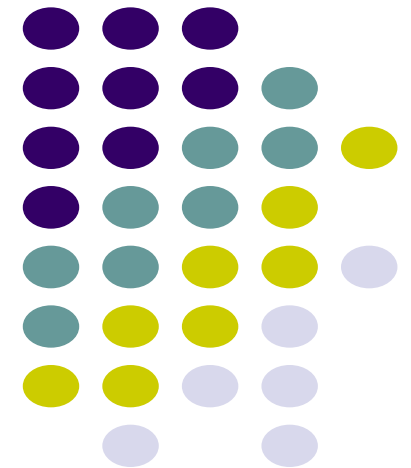
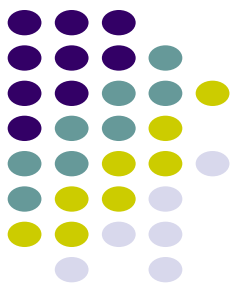


Arquitetura e Organização de Computadores

Armazenamento e E/S Parte I

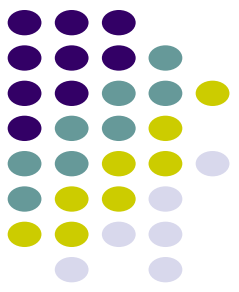
Prof. Sílvio Fernandes





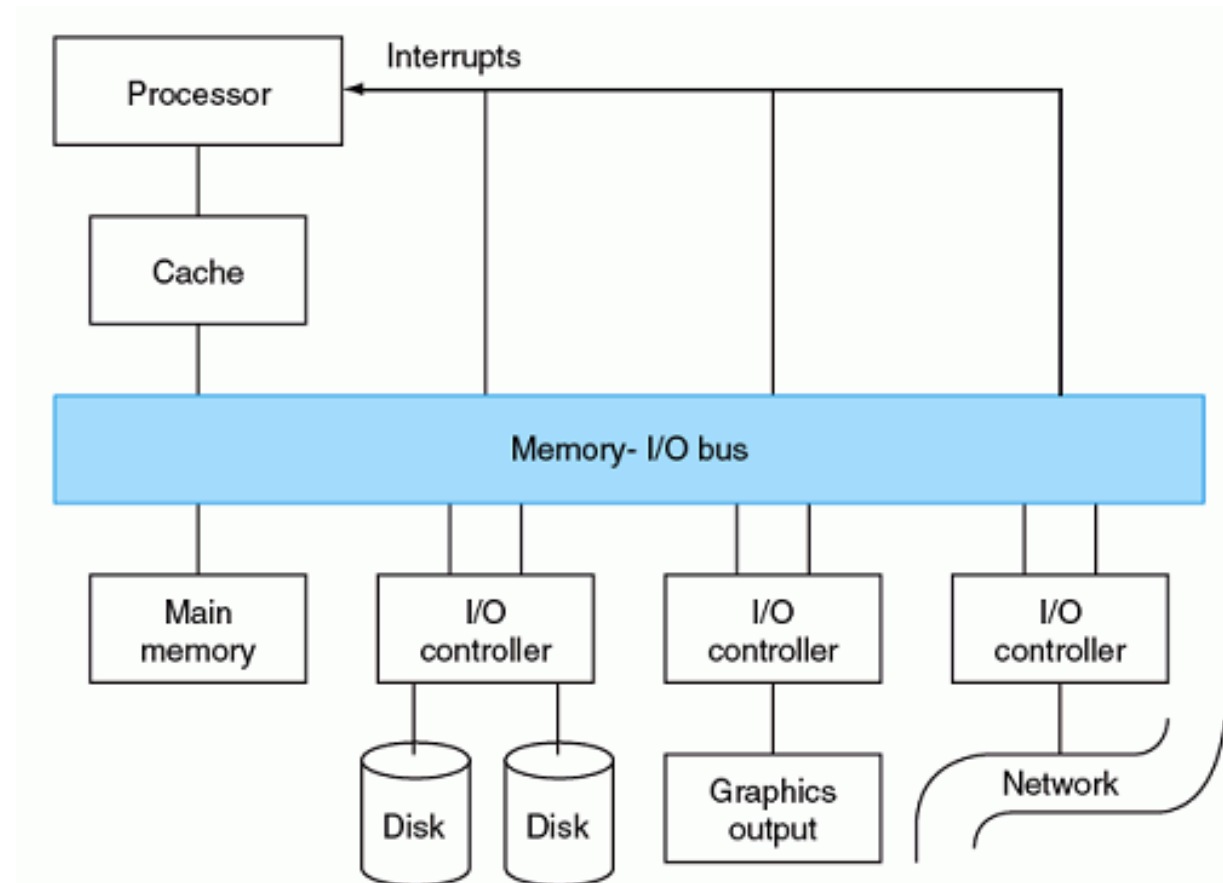
Introdução

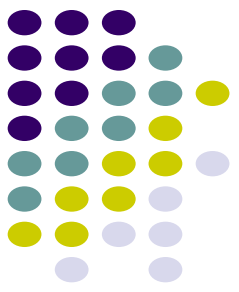
- Usuários se frustram com computadores travando
- Mas eles não tolerariam se o sistema de armazenamento falhasse
 - Tolerância à confiabilidade é muito mais alta em relação ao armazenamento do que à computação
 - Os sistemas de E/S também precisam planejar a facilidade de expansão e a diversidade
 - Desempenho tem um papel secundário
 - Depende de muitos aspectos: características dos dispositivos, da conexão entre dispositivo e o resto do sistema, hierarquia de memória e do sist. operacional



Introdução

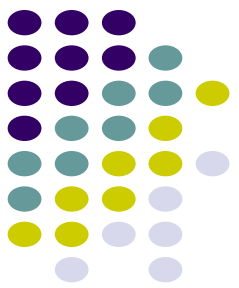
- Estrutura simples de um sistema de E/S





Introdução

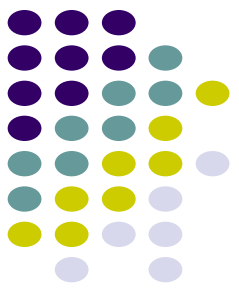
- Características da E/S
 - **Comportamento:** entrada (somente entrada), saída (somente escrita, não pode ser lido) ou armazenamento (pode ser lido e normalmente reescrito)
 - **Parceria:** um humano ou uma máquina está na outra extremidade do dispositivo de E/S, seja alimentando ou lendo os dados
 - **Taxa de dados:** a taxa de pico em que os dados podem ser transferidos entre o dispositivo de E/S e a memória principal



Introdução

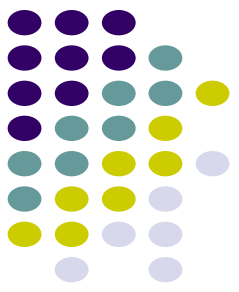
- A diversidade de dispositivos de E/S

Device	Behavior	Partner	Data rate (Mbit/sec)
Keyboard	Input	Human	0.0001
Mouse	Input	Human	0.0038
Voice input	Input	Human	0.2640
Sound input	Input	Machine	3.0000
Scanner	Input	Human	3.2000
Voice output	Output	Human	0.2640
Sound output	Output	Human	8.0000
Laser printer	Output	Human	3.2000
Graphics display	Output	Human	800.0000–8000.0000
Cable modem	Input or output	Machine	0.1280–6.0000
Network/LAN	Input or output	Machine	100.0000–10000.0000
Network/wireless LAN	Input or output	Machine	11.0000–54.0000
Optical disk	Storage	Machine	80.0000–220.0000
Magnetic tape	Storage	Machine	5.0000–120.0000
Flash memory	Storage	Machine	32.0000–200.0000
Magnetic disk	Storage	Machine	800.0000–3000.0000



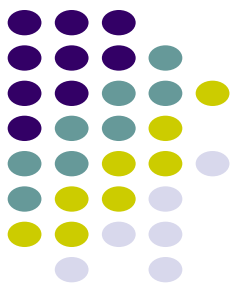
Introdução

- Avaliar o desempenho da E/S depende da aplicação
 - Quantos dados podemos mover pelo sistema em determinado momento?
 - Quantas operações de E/S podemos realizar por unidade de tempo?
 - Em algumas aplicações a preocupação pode ser a vazão e facilidade de expansão (servidores)
 - Em outra o tempo de resposta e diversidade dos dispositivos (desktops e sistemas embarcados)



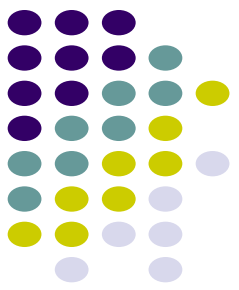
Confiança, confiabilidade e disponibilidade

- Armazenamento confiável (Laprie, 1985):
 - *Confiança de um sistema computacional é a qualidade do serviço entregue de modo que a confiança passa a ser justificadamente depositada sobre esse serviço. O serviço entregue por um sistema é o seu comportamento real observado como percebido por outro(s) sistema(s) interagindo com os usuário desse sistema. Cada módulo possui um comportamento especificado ideal, no qual uma especificação de serviço é uma descrição combinada do comportamento esperado. Uma falha do sistema ocorre quando o comportamento real se desvia do comportamento especificado*



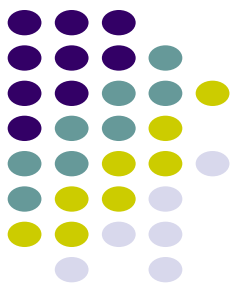
Confiança, confiabilidade e disponibilidade

- Usuários podem ver um sistema alternar entre 2 estados
 1. **Realização do serviço:** entregue conforme especificado
 2. **Interrupção do serviço:** entregue diferente do serviço especificado
 - Transições de 1 para 2: falhas
 - Transições de 2 para 1: restaurações
- A oscilação desses estados define: confiabilidade e disponibilidade



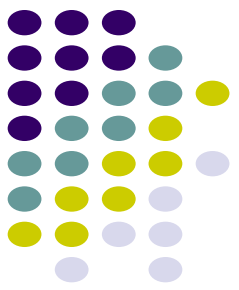
Confiança, confiabilidade e disponibilidade

- Confiabilidade
 - Medida da realização contínua do serviço (tempo para falha)
 - *Tempo médio para falha dos discos (MTTF)* é uma medida de confiabilidade
 - *Taxa de falha anual (AFR)* é simplesmente a porcentagem dos dispositivos falhariam em um ano para determinado MTTF
 - A interrupção do serviço é medida como *tempo médio para reparo (MTTR)*
 - *Tempo médio entre falhas (MTBF)* é a soma $MTTF + MTTR$



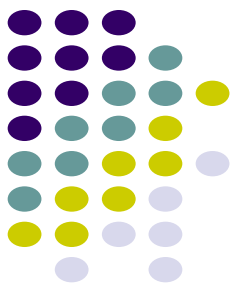
Confiança, confiabilidade e disponibilidade

- Disponibilidade
 - Realização do serviço com relação à alternância entre os 2 estados de realização e interrupção
 - $\text{Disponibilidade} = \text{MTTF} / (\text{MTTF} + \text{MTTR})$
- Confiabilidade e disponibilidade são medida quantificáveis, e não apenas sinônimo de confiança



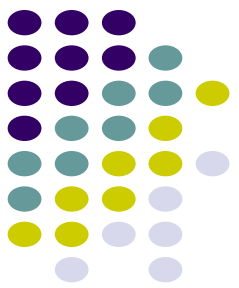
Confiança, confiabilidade e disponibilidade

- Três maneiras de melhorar o MTTF
 1. **Impedimento de falha:** evitar a ocorrência da falha pela construção
 2. **Tolerância a falhas:** uso de redundância para permitir que o serviço cumpra com a especificação de serviço apesar da ocorrência de falhas
 3. **Previsão de falhas:** prevê a presença e criação de falhas, o que se aplica a falhas do HW e do SW, permitindo que o componente seja substituído antes de falhar
- Diminuir o MTTR pode ajudar na disponibilidade quanto aumentar o MTTF



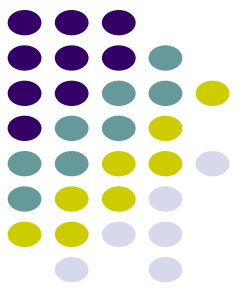
Armazenamento em disco

- Os discos contam com pratos giratório cobertos por uma superfície magnética
- Utiliza uma cabeça de leitura/escrita móvel para acessar o disco
- O armazenamento é **não volátil**
- O disco consiste em uma coleção de pratos (1-4) que giram entre 5.400 a 15.000 RPM e tem diâmetro entre 2,5 cm e 9 cm
- Cada superfície está dividida em círculos concêntricos (**trilhas**) as quais são divididas em **setores**
- Normalmente existem de 10.000 a 50.000 trilhas e cada trilha pode ter de 100 a 500 setores (de 512 bytes a 4.096 bytes)



Armazenamento em disco

- O termo **cilindro** é usado para se referir a todas as trilhas sob as cabeças em determinado ponto para todas as superfícies
- Para acessar os dados
 1. Posicionar a cabeça sobre a trilha apropriada (**seek**)
 2. Esperar até o setor desejado girar sob a cabeça de leitura/escrita (**latência rotacional** ou **atraso**)
 3. Transferir os dados
- Os fabricantes costumam informar o tempo de seek mínimo, máximo e médio
- O tempo de seek real pode ser de apenas 25% a 33% do número anunciado



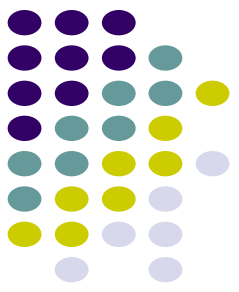
Armazenamento em disco

- A latência rotacional está a meio caminho ao redor do disco
- Para discos que giram entre 5.400 e 15.000 RPM, a latência rotacional média é

$$\begin{aligned}\text{Average rotational latency} &= \frac{0.5 \text{ rotation}}{5400 \text{ RPM}} = \frac{0.5 \text{ rotation}}{5400 \text{ RPM} / \left(60 \frac{\text{seconds}}{\text{minute}}\right)} \\ &= 0.0056 \text{ seconds} = 5.6 \text{ ms}\end{aligned}$$

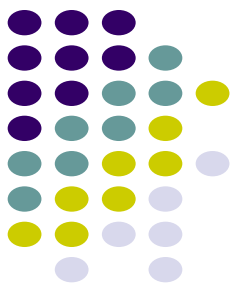
and

$$\begin{aligned}\text{Average rotational latency} &= \frac{0.5 \text{ rotation}}{15,000 \text{ RPM}} = \frac{0.5 \text{ rotation}}{15,000 \text{ RPM} / \left(60 \frac{\text{seconds}}{\text{minute}}\right)} \\ &= 0.0020 \text{ seconds} = 2.0 \text{ ms}\end{aligned}$$



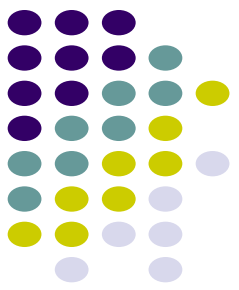
Armazenamento em disco

- O *tempo de transferência* é uma função do tamanho do setor, da velocidade de rotação e da densidade de gravação de uma trilha
- As taxa de transferência em 2008 estavam entre 70 e 125 MB/s
- As controladores de discos possuem caches internas e suas taxas de transferências chegam a 375 MB/s (3 Gbit/s)



Armazenamento em disco

- **Exemplo:** qual é o tempo médio para ler ou escrever um setor de 512 bytes em um disco típico girando a 15.000 RPM? O tempo de seek médio anunciado é de 4ms, a taxa de transferência é de 100MB/s e o overhead da controladora é de 0,2ms. Suponha que o disco esteja ocioso, de modo que não existe um tempo de espera.



Armazenamento em disco

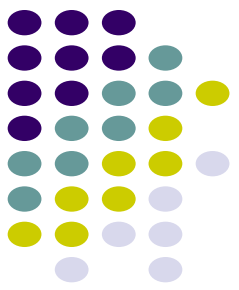
- **Exemplo:**

- O tempo médio de acesso ao disco é igual ao Tempo médio de seek + Atraso rotacional médio + Tempo de transferência + Overhead da controladora

$$4,0\text{ms} + (0,5/15.000 \text{ RPM}) + (0,5 \text{ KB}/100 \text{ MB/s}) + 0,2\text{ms} =$$

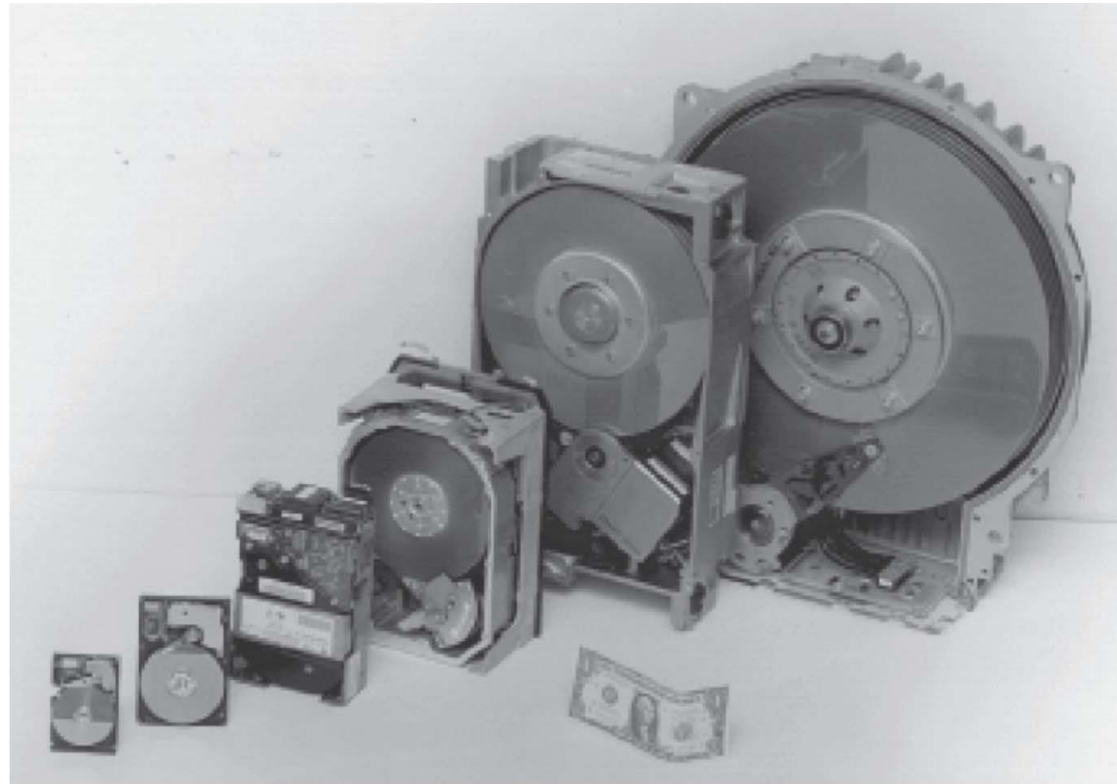
$$4,0 + 2,0 + 0,005 + 0,2 = 6,2 \text{ ms}$$

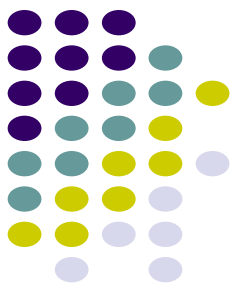
- Se o tempo médio de seek for 25% do anunciado:
 - $1,0\text{ms} + 2,0\text{ms} + 0,005\text{ms} + 0,2\text{ms} = 3,2\text{ms}$



Armazenamento em disco

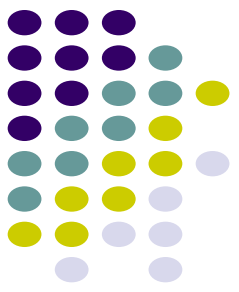
- A densidades de disco têm continuado a aumentar há mais de 50 anos
 - Impacto no tamanho e preço





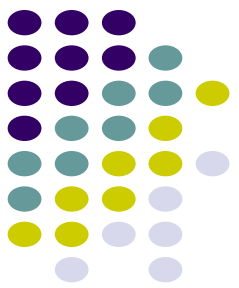
Armazenamento em disco

Characteristics	Seagate ST33000655SS	Seagate ST31000340NS	Seagate ST973451SS	Seagate ST9160821AS
Disk diameter (inches)	3.50	3.50	2.50	2.50
Formatted data capacity (GB)	147	1000	73	160
Number of disk surfaces (heads)	2	4	2	2
Rotation speed (RPM)	15,000	7200	15,000	5400
Internal disk cache size (MB)	16	32	16	8
External interface, bandwidth (MB/sec)	SAS, 375	SATA, 375	SAS, 375	SATA, 150
Sustained transfer rate (MB/sec)	73–125	105	79–112	44
Minimum seek (read/write) (ms)	0.2/0.4	0.8/1.0	0.2/0.4	1.5/2.0
Average seek read/write (ms)	3.5/4.0	8.5/9.5	2.9/3.3	12.5/13.0
Mean time to failure (MTTF) (hours)	1,400,000 @ 25°C	1,200,000 @ 25°C	1,600,000 @ 25°C	—
Annual failure rate (AFR) (percent)	0.62%	0.73%	0.55%	—
Contact start-stop cycles	—	50,000	—	>600,000
Warranty (years)	5	5	5	5
Nonrecoverable read errors per bits read	<1 sector per 10 ¹⁶	<1 sector per 10 ¹⁵	<1 sector per 10 ¹⁶	<1 sector per 10 ¹⁴
Temperature, shock (operating)	5°–55°C, 60 G	5°–55°C, 63 G	5°–55°C, 60 G	0°–60°C, 350 G
Size: dimensions (in.), weight (pounds)	1.0" × 4.0" × 5.8", 1.5 lbs	1.0" × 4.0" × 5.8", 1.4 lbs	0.6" × 2.8" × 3.9", 0.5 lbs	0.4" × 2.8" × 3.9", 0.2 lbs
Power: operating/idle/standby (watts)	15/11/—	11/8/1	8/5.8/—	1.9/0.6/0.2
GB/cu. in., GB/watt	6 GB/cu.in., 10 GB/W	43 GB/cu.in., 91 GB/W	11 GB/cu.in., 9 GB/W	37 GB/cu.in., 84 GB/W
Price in 2008, \$/GB	~ \$250, ~ \$1.70/GB	~ \$275, ~ \$0.30/GB	~ \$350, ~ \$5.00/GB	~ \$100, ~ \$0.60/GB



Armazenamento flash

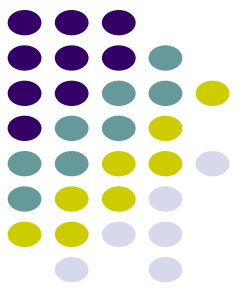
- O 1º. desafiador dos discos convincente
- Memória semicondutora somente de leitura programável eletricamente (EEPROM)
- Não volátil com latência 100-1000 vezes mais rápida que o disco, menor, gasta menos energia e é mais resistente ao choque
- O custo por gigabyte ainda é mais alto que os discos



Armazenamento flash

- O preço da flash tem caído

Characteristics	Kingston SecureDigital (SD) SD4/8 GB	Transend Type I CompactFlash TS16GCF133	RiDATA Solid State Disk 2.5 inch SATA
Formatted data capacity (GB)	8	16	32
Bytes per sector	512	512	512
Data transfer rate (read/write MB/sec)	4	20/18	68/50
Power operating/standby (W)	0.66/0.15	0.66/0.15	2.1/—
Size: height × width × depth (inches)	0.94 × 1.26 × 0.08	1.43 × 1.68 × 0.13	0.35 × 2.75 × 4.00
Weight in grams (454 grams/pound)	2.5	11.4	52
Mean time between failures (hours)	> 1,000,000	> 1,000,000	> 4,000,000
GB/cu. in., GB/watt	84 GB/cu.in., 12 GB/W	51 GB/cu.in., 24 GB/W	8 GB/cu.in., 16 GB/W
Best price (2008)	~ \$30	~ \$70	~ \$300



Armazenamento flash

- Existem memórias flash NOR e NAND

Characteristics	NOR Flash Memory	NAND Flash Memory
Typical use	BIOS memory	USB key
Minimum access size (bytes)	512 bytes	2048 bytes
Read time (microseconds)	0.08	25
Write time (microseconds)	10.00	1500 to erase + 250
Read bandwidth (MBytes/second)	10	40
Write bandwidth (MBytes/second)	0.4	8
Wearout (writes per cell)	100,000	10,000 to 100,000
Best price/GB (2008)	\$65	\$4

- Na flash os bits se desgastam

Referências

- PATTERSON, D. A. ; HENNESSY, J.L. Organização e projeto de computadores – a interface hardware software. 4. ed. Editora Campus, 2014.

