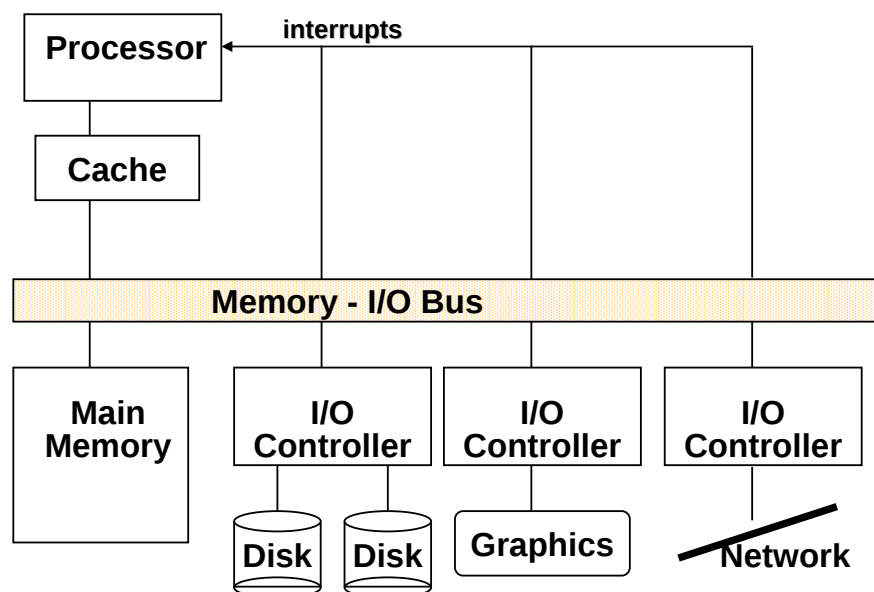


Aula 17: I/O — Dispositivos de Armazenamento, Métricas, e Produtividade

Subsistemas de I/O



$$\text{Time(workload)} = \text{Time(CPU)} + \text{Time(I/O)} - \text{Time(Overlap)}$$



Capítulo 6

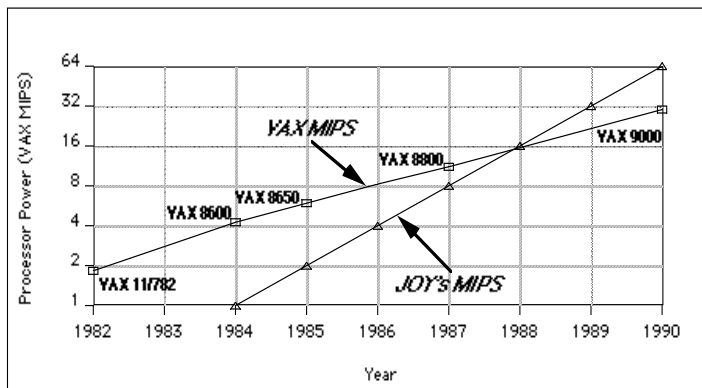
- Contexto histórico de I/O
- Dispositivos de armazenamento secundário e terciário
- Medidas de performance de I/O
- Intro. a teoria das filas
- Interface do processador e barramentos de I/O
- *Redundant Array of Inexpensive Disks* (RAID)
- Sistema de arquivo do Unix
- Benchmarks de I/O

Motivação: Por que estudar I/O?

- Performance de CPUs: 50% a 100% por ano
- Supercomputadores c/ multiprocessadores: 150% por ano
- Performance de subsistema de I/O system é limitada por atrasos mecânicos
< 5% por ano (IO por seg ou MB por seg)
- Lei de Amdahl: Speedup limitado por parte mais lenta
 - 10% IO & 10x CPU => 5x (Perde 50%)
 - 10% IO & 100x CPU => 10x (Perde 90%)
- Gargalo de I/O:
CPU estão ficando mais rápidas
Velocidade de CPUs está fazendo menos diferença

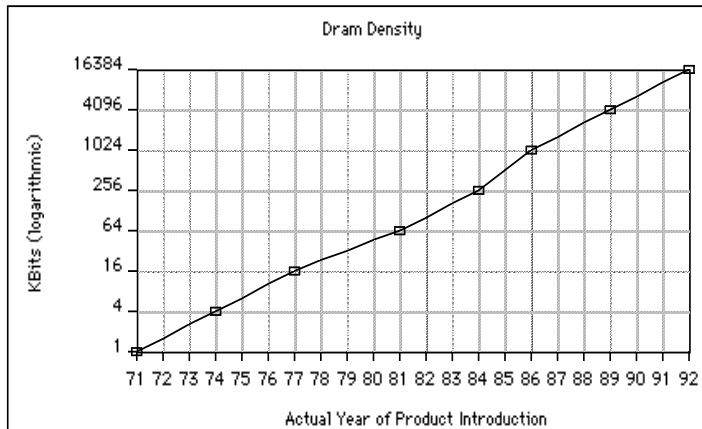


Avanços Tecnológicos



Performance de CPU

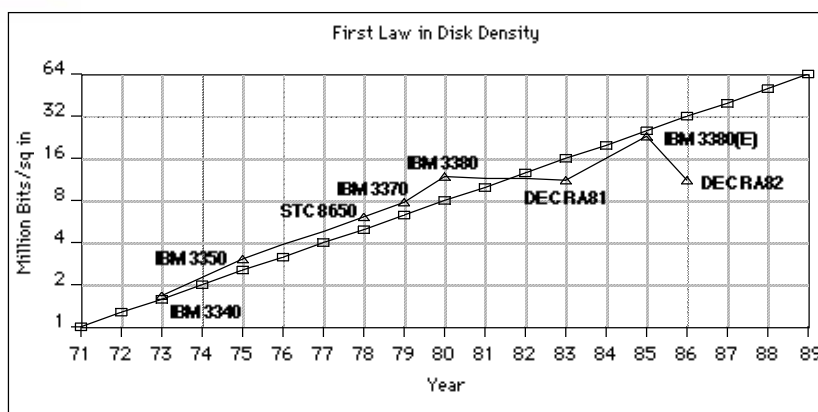
- Mini:
40% / ano
- RISC:
100% / ano



Capacidade de DRAM 2x a cada 2-3 anos



Avanços Tecnológicos



Capacidade de Discos 2x a cada 3 anos

- Hoje: Poder de processamento dobra a cada 18 meses
- Hoje: Tamanho de memória dobra a cada 18 meses (?)
- Hoje: Capacidade de disco dobra a cada 18 meses
- Velocidade de posicionamento (Seek + Rotate) dobra a cada 10 anos!

I/O vs o resto

Dispositivos de Armazenamento

- Perspectiva
 - 1950s: migração de processamento em batch p/ on-line
 - 1990s: migração para dispositivos únicos
 - computadores em telefones, livros eletrônicos, carros, câmeras de vídeo, ...
 - redes de alta velocidade
- Efeitos na indústria:
 - Armazenamento embutido
 - menor, mais barato, mais confiável, baixa potência
 - Utilitários
 - gerenciamento de dispositivos de alta capacidade, hierarquizado

Perspectiva Histórica

- 1956 IBM Ramac — predecessor do Winchester
 - desenvolvido para mainframes
 - interfaces proprietárias
- 1970s
 - floppy de 5.25
 - memória semicondutora e microprocessadores
 - primeiras interfaces de disco
 - ST506, SASI, SMD, ESDI

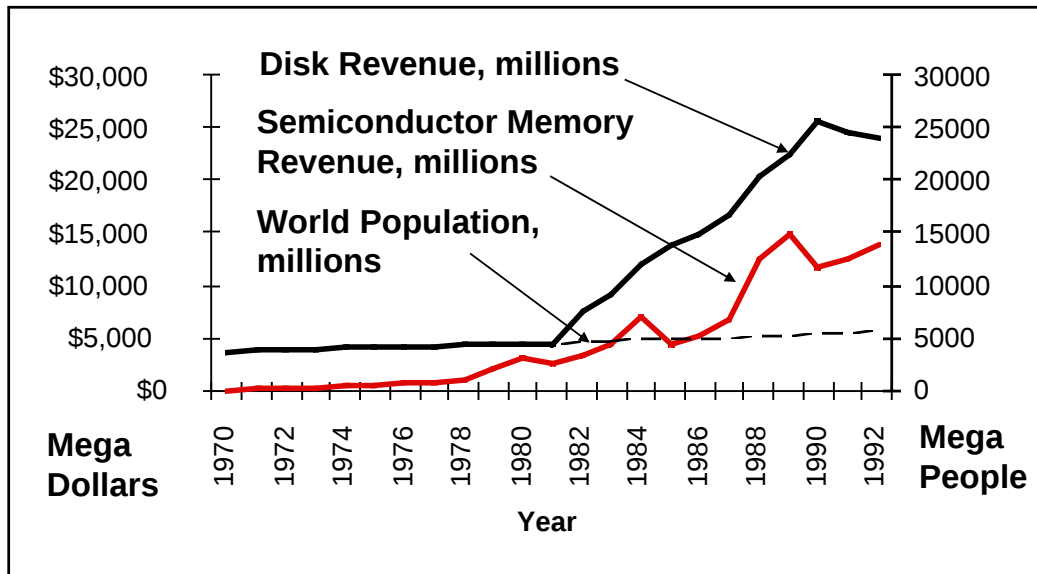
Perspectivas Históricas

- Início dos anos 1980s
 - PCs e workstations
- Meio dos anos 1980s
 - Cliente/servidor
 - Servidor de arquivos
 - 8 polegadas para 5.25 polegadas
 - Disco para as massas
 - padrões da indústria: SCSI, IPI, IDE
 - discos de 5.25'' para PCs
 - fim das interfaces proprietárias

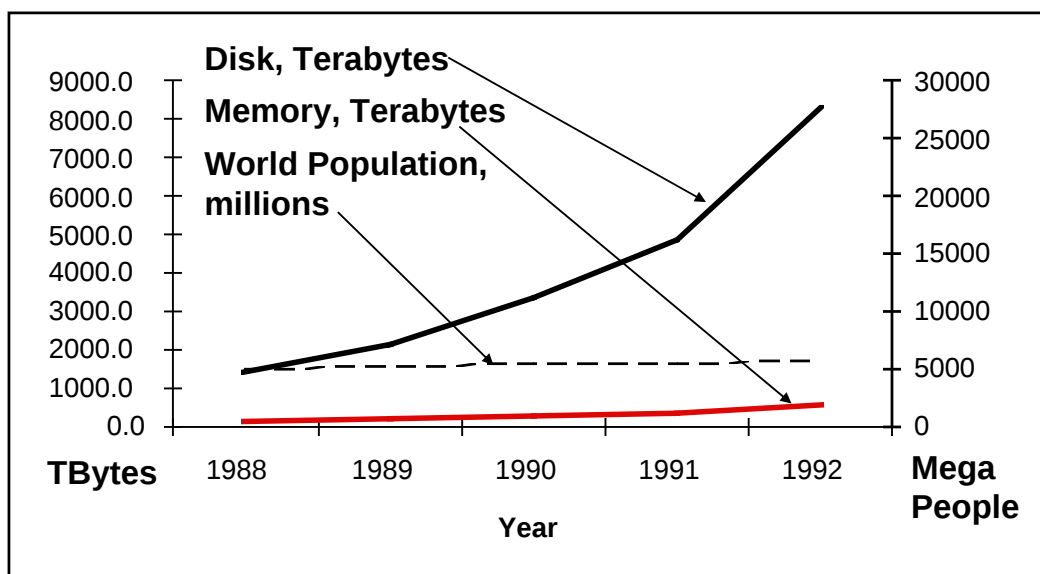
Perspectivas Históricas

- Fim dos 1980s/Início dos 1990s:
 - Laptops, notebooks, palmtops
 - discos de 3.5'', 2.5'', 1.8'', 1.3''
 - Tamanho e capacidade dita mercado de discos (não é performance)
 - RAM e flash RAM em PCMCIAs
 - caras
 - Performance de discos óticos é baixa como dispositivo secundário (CD ROM)

Perspectiva Histórica



Perspectiva Histórica



1.5 MBytes por pessoa na terra vendido em 1992
0.1 MBytes por pessoa na terra vendido em 1992



Tecnologias Alternativas para Armazenamento



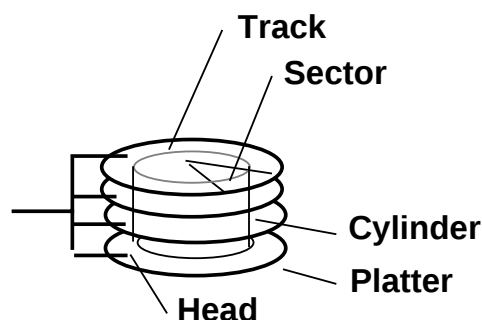
Technology	Cap (MB)	BPI	TPI	BPI*TPI (Million)	Data Xfer (KByte/s)	Access Time
Conventional Tape:						
Cartridge (.25")	150	12000	104	1.2	92	minutes
IBM 3490 (.5")	800	22860	38	0.9	3000	seconds
Helical Scan Tape:						
Video (8mm)	4600	43200	1638	71	492	45 secs
DAT (4mm)	1300	61000	1870	114	183	20 secs
D-3 (1/2")	20,000					15 secs?
Magnetic & Optical Disk:						
Hard Disk (5.25")	1200	33528	1880	63	3000	18 ms
IBM 3390 (10.5")	3800	27940	2235	62	4250	20 ms
Sony MO (5.25")	640	24130	18796	454	88	100 ms



Dispositivos: Discos Magnéticos



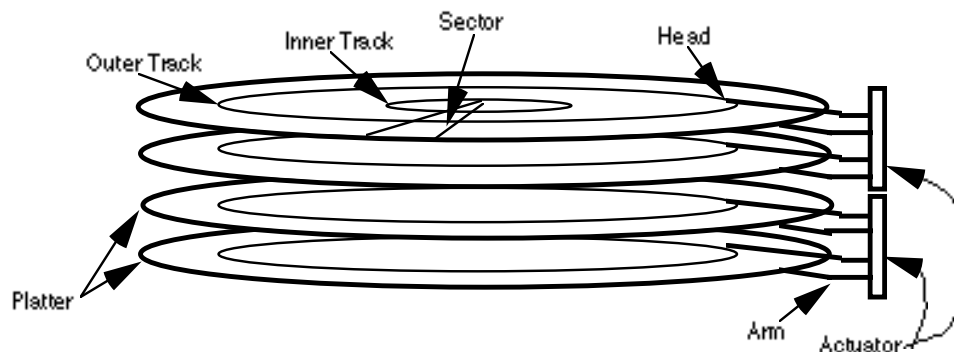
- **Objetivos:**
 - Armazenamento não-volátil por tempo longo
 - Alta capacidade
- **Características:**
 - Tempo de acesso alto
 - Seek time, rotational delay
 - Transfer rate (1 setor/ms)
- **Capacidade**
 - Gigabytes



3600 RPM = 60 RPS => 16 ms per rev
 ave rot. latency = 8 ms
 32 sectors per track => 0.5 ms per sector
 1 KB per sector => 2 MB / s
 32 KB per track
 20 tracks per cyl => 640 KB per cyl
 2000 cyl => 1.2 GB

Tempo de resposta = Fila + Controlador + Seek + Rot + Xfer

Terminologia



Disk Latency = Queuing Time + Seek Time + Rotation Time + Xfer Time

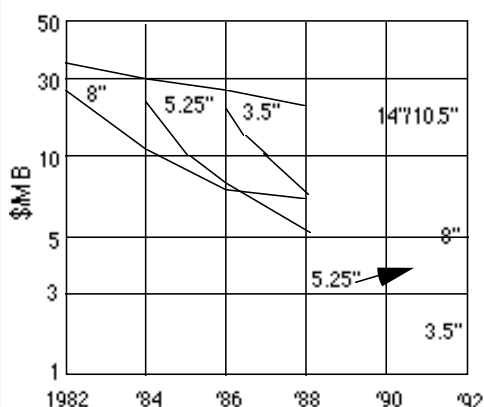
Ordem de magnitude para transferências de 4Kb:

Seek: 15 ms ou menos

Rotation: 8.3 ms @ 3600 rpm (4.2 ms @ 7200 rpm)

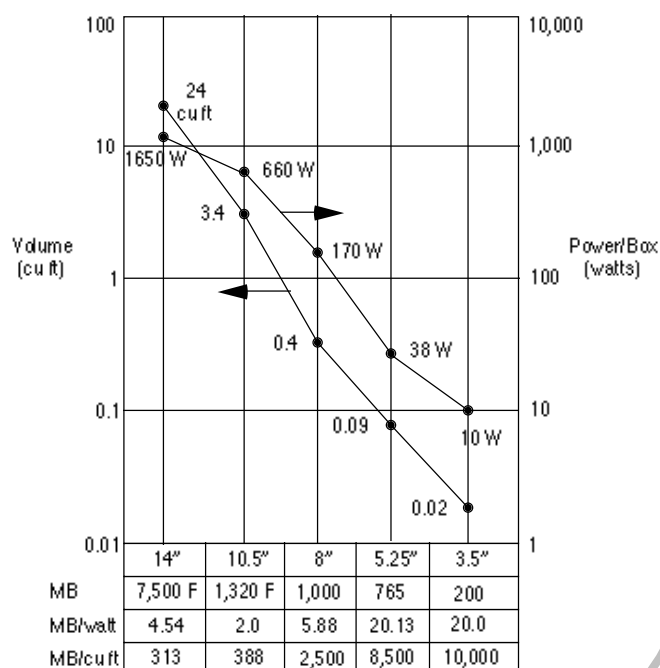
Xfer: 2 ms @ 3600 rpm (1 ms @ 7200 rpm)

Vantagens de Discos Pequenos



**Baixo custo/MB
Alto MB/volume
Alto MB/watt
Baixo custo/Atuador**

Custo e eficiências ambientais



Tape vs. Disco

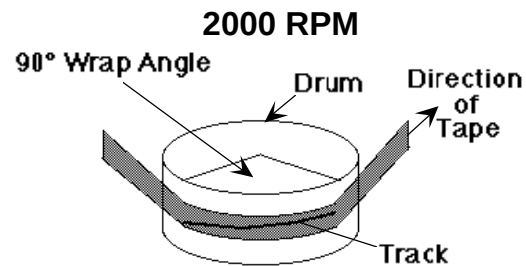
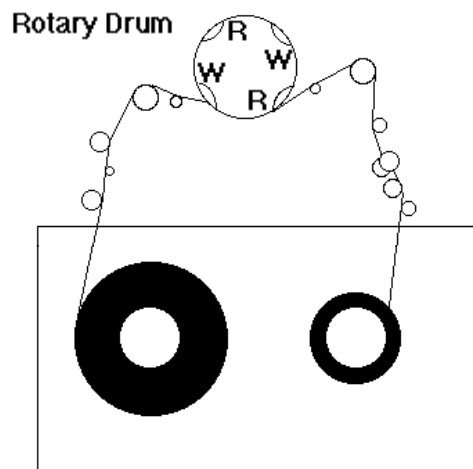
- Tape utiliza mesma tecnologia que disco para densidades longitudinais
- Custo performance baseado em tecnologia:
 - pratos giratórios
 - (acesso aleatório, área limitada)
 - fitas longas de material magnético
 - (acesso seqüencial, tamanho ilimitado)

Exemplo: R-DAT

Cabeça rotativa para *Digital Audio Tape*

- Maior densidade de gravação disponível atualmente
- Melhoria da ordem de 10X até 1999
 - Velocidade: melhoria da velocidade do tape e do tambor
 - Capacidade: melhoria na densidade

Tecnologia

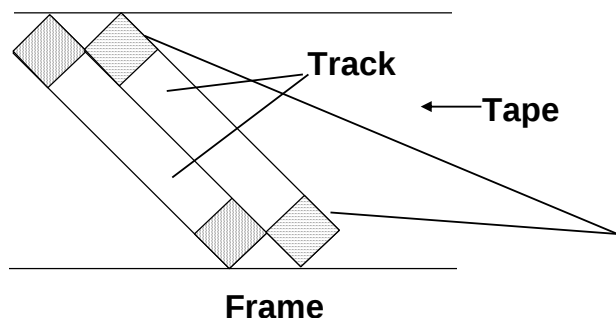


Quatro cabeças para gravação

Verificação de gravação por leitura

Tecnologia R-DAT

DDS ANSI Standard (HP, SONY)



65% da trilha é para dados
70% Bytes de Dados
30% Bytes de Paridade +
Códigos Reed-Solomon

Trilha (2900 Bytes de dados)

Quadro (2 trilhas)

Grupo (22 quadros + grupo opcional ECC)

Probabilidade teórica de erros:

- sem ECC: 1 em 10^{26}
- com ECC: 1 em 10^{33}



Discos Óticos vs. Tape

	Discos Óticos	Tapes
Tipo	5.25"	8mm
Capacidade	0.75 GB	5 GB
\$ mídia	\$90 - \$175	\$8
\$ drive	\$3,000	\$3,000
acesso	Write Once	Read/Write



Desvantagens de Tape

- Desgaste do tape:
 - 100s a 1000s
- Desgaste da cabeça:
 - 2000 horas
- Incluídos no modelo de confiabilidade
- Tempo longo para *rewind, eject, load*



Performance Relativa dos Dispositivos - 1995

Discos Magnéticos

5.25"	9.1 GB	\$2129	\$0.23/MB
3.5"	4.3 GB	\$1199	\$0.27/MB
2.5"	514 MB	\$299	\$0.58/MB

Discos Óticos

5.25"	4.6 GB	\$1695+199	\$0.41/MB
-------	--------	------------	-----------

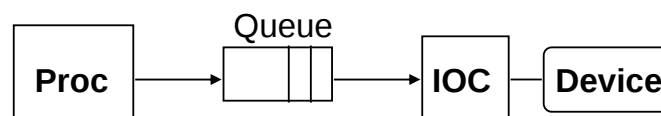
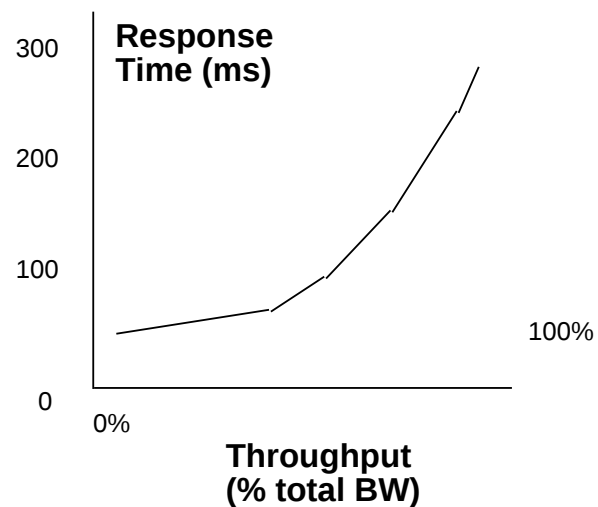
Cartões PCMCIA

RAM estática	4.0 MB	\$700	\$175/MB
Flash RAM	40.0 MB	\$1300	\$32/MB
	175 MB	\$3600	\$20.50/MB



Performance do Subsystema de I/O de Disco

Métricas:
Tempo de resposta
Throughput



Response time = Queue + Device Service time

Tempo de Resposta vs. Produtividade

- Ambientes interativos:

Cada *transação* possui 3 partes:

- *Tempo de entrada* : usuário entra com o comando
- *Tempo de resposta do sistema*: sistema responde
- *Tempo de análise*: tempo até nova requisição de comando

transação 1



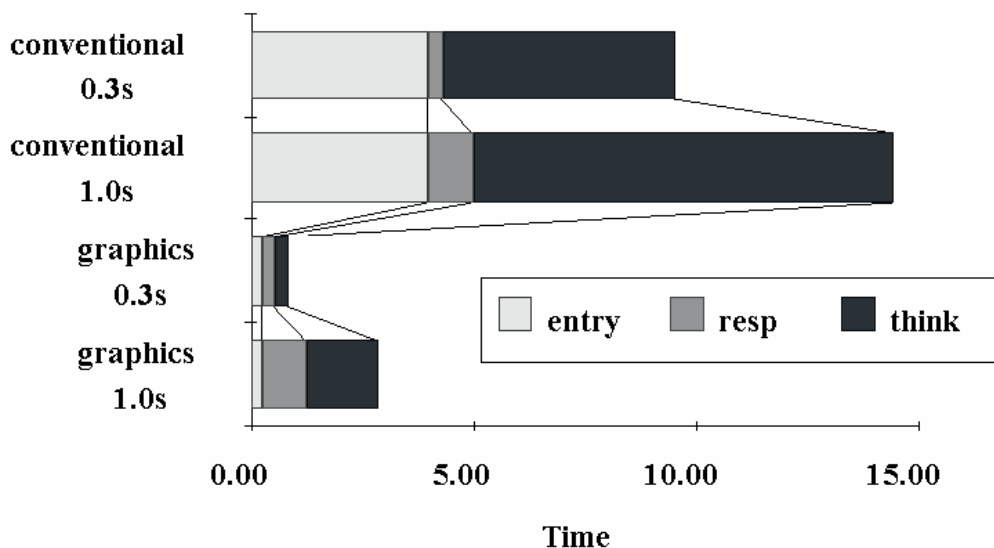
transação 2



- O que acontece com tempo de transação se tempo de resposta do sistema diminui de 1.0 seg para 0.3 seg?

- c/ teclado: 4.0 seg. entrada, 9.4 seg. análise
- c/ gráficos: 0.25 seg. entrada, 1.6 seg. análise

Tempo de Resposta vs. Produtividade



- Aumento de produtividade (34% a 70%)
- Outro estudo: novato com sistema mais rápido = especialista com sistema mais lento