

Ch.11 – I/O Management and Disk Scheduling

11.1 – Input / Output Devices

11.2 – Organization of The I/O Function

11.2.1 - The Evolution of the I/O Function

11.2.2 - Direct Memory Access

11.3 – Operating System Design Issues

11.3.1 - Design Objectives

11.3.2 - Logical Structures of the I/O Function

11.4 – Input / Output Buffering

... Ch.11 – I/O Management and Disk Scheduling

11.5 – Disk Scheduling

11.5.1 - Disk Performance Parameters

11.5.2 - Disk Scheduling Policies

11.6 – Redundant Array I Disk

11.7 – Disk Cache

11.7.1 - Design Considerations

11.7.2 - Performance Consideration

... Ch.11 – I/O Management and Disk Scheduling

11.8 – UNIX SVR4 Input / Output (LEITURA COMPLEMENTAR)

11.9 – LINUX Input / Output (LEITURA COMPLEMENTAR)

11.10 – WINDOWS Input / Output (LEITURA COMPLEMENTAR)

Referências Bibliográficas

- Operating Systems – Internals and Design Principles. William Stallings. 7th, Prentice-Hall 2012.
- Instructor Resources – Operating Systems - 7th
<http://williamstallings.com/OperatingSystems/OS7e-Instructor/>

An artifact can be thought of as a meeting point - an “interface” in today’s terms between an “inner” environment, the substance and organization of the artifact itself, and an “outer” environment, the surroundings in which it operates. If the inner environment is appropriate to the outer environment, or vice versa, the artifact will serve its intended purpose.

The Sciences of The Artificial, Herbert Simon

11. I/O Management / 11.1. Input/Output Devices

11.1 – I/O Deives

- aaa

11. I/O Management / 11.2. Organization of the I/O Function

11.2 – Organization of the I/O Function

- aaa

11. I/O Management / 11.2. Organization of the I/O Function

11.2.1 – The Evolution of the I/O Function

- aaa

11. I/O Management / 11.2. Organization of the I/O Function

11.2.2 – Direct Memory Access

- aaa

11. I/O Management / 11.3. Operation System Design Issues

11.3 – Operation System Design Issues

- aaa

11. I/O Management / 11.3. Operation System Design Issues

11.3.1 – Design Objectives

- aaa

11. I/O Management / 11.3. Operation System Design Issues

11.3.2 – Logical Structures of the I/O Function

- aaa

11. Input / Output Management / 11.4. Input/Output Buffering

11.4 – Input/Output Buffering

- aaa

11. Input / Output Management / 11.4. Input/Output Buffering

11.4.1 – Single Buffer

- aaa

11. Input / Output Management / 11.4. Input/Output Buffering

11.4.2 – Double Buffer

- aaa

11. Input / Output Management / 11.4. Input/Output Buffering

11.4.3 – Circular Buffer

- aaa

11. Input / Output Management / 11.4. Input/Output Buffering

11.4.4 – The Utility of Buffering

- aaa

11.5 – Disk Scheduling

- “**last 40 years**” .. aumento na velocidade dos processadores e da memória principal superou em muito o do acesso ao disco.
- .. com as velocidades do processador e da memória principal aumentando em cerca de 02 ordens de magnitude em comparação com 01 ordem de magnitude para o disco.
- “**resultado**” .. discos são atualmente pelo menos 04 ordens de magnitude mais lentos que a memória principal.
- .. assim o desempenho do subsistema de armazenamento em disco é uma preocupação vital e que evidencia a necessidade de novos esquemas para melhorar esse desempenho.

11.5.1 – Disk Performance Parameters

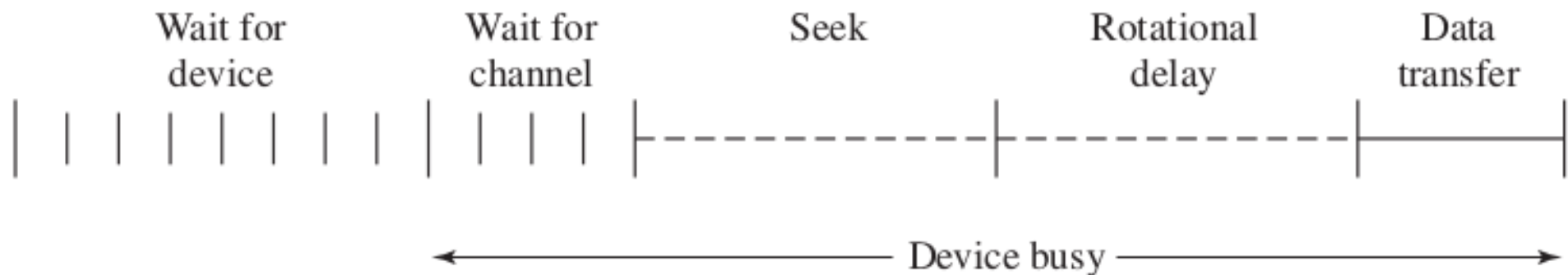
- **“disk input/output operation”** .. depende do sistema do computacional, do sistema operacional e da natureza do canal de I/O e do hardware do controlador de disco.
- **“read / write operation”** .. para ler ou escrever em um disco que gira a velocidade constante, o cabeçote deve estar posicionado na trilha desejada e no início do setor desejado nessa trilha.
- **“seek time”** .. em um sistema com cabeçote móvel, o tempo para posicionar o cabeçote na trilha é conhecido como tempo de busca.



11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.1 – Disk Performance Parameters

- “**rotational delay**” .. tempo que leva para o início do setor atingir a cabeça é conhecido como atraso rotacional ou latência rotacional.
- “**access time**” .. soma do tempo de busca, se houver, e o atraso rotacional é igual o tempo de acesso, que é o tempo que leva para se posicionar para ler ou escrever.
- “**transfer time**” .. operação de leitura ou escrita é então executada à medida que o setor se move sob o cabeçote, tempo no qual há a transferência de dados da operação;



11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.1 – Disk Performance Parameters

- “**several queueing delays**” .. além do tempo de acesso e do tempo de transferência, existem vários atrasos de fila normalmente associados a uma operação de I/O de disco.
- .. quando um processo emite uma solicitação de I/O, ele deve primeiro aguardar em uma fila para que o dispositivo esteja disponível e, nesse momento, o dispositivo é atribuído ao processo.
- .. se o dispositivo compartilhar um único canal de I/O ou um conjunto de canais de I/O com outras unidades de disco, pode ser necessário aguardar que o canal esteja disponível.
- .. nesse ponto, a busca é executada para iniciar o acesso ao disco.

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.1 – Disk Performance Parameters

- **“Rotational Positional Sensing”** .. em alguns sistemas avançados para servidores, usa-se uma técnica conhecida como sensoramento de posição rotacional (RPS).
- .. quando o comando de busca é emitido, o canal é liberado para lidar com outras operações de I/O e quando a busca é concluída, o dispositivo determina quando os dados irão girar sob a cabeça.
- .. à medida que esse setor se aproxima da cabeça, o dispositivo tenta restabelecer o caminho de comunicação de volta ao “host”.
- .. se a unidade de controle ou o canal estiver ocupado com outra I/O, a tentativa de reconexão falha e o dispositivo deve girar uma volta inteira antes de tentar reconectar, o que é chamado de falha de RPS.
- .. este é um elemento de atraso extra que deve ser adicionado à linha do tempo – **“seek time”**.

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.1 – Disk Performance Parameters

- “**seek time**” .. tempo de busca é o tempo necessário para mover o braço do disco para a trilha desejada, no entanto, é uma medida difícil de ser medida face a posição do cabeçote.
- .. tempo de busca consiste em 02 componentes, ou seja, tempo inicial de inicialização e o tempo necessário para percorrer as trilhas que devem ser cruzadas uma vez que o braço de acesso esteja atualizado.
- .. infelizmente, o tempo de travessia não é uma função linear do nro. de trilhas, mas inclui um ajuste após posicionar a cabeça sobre a trilha de destino até que a identificação da pista seja confirmada.
- e.g., há alguns anos, um disco típico tinha 36 cm de diâmetro, enquanto que hoje o tamanho comum é de 8,9 cm, o que reduz a distância que o braço precisa percorrer » “seek time” = 10 ms.

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.1 – Disk Performance Parameters

- “**rotational delay**” .. tempo necessário para que a área endereçada do disco gire para uma posição onde seja acessível pelo cabeçote de leitura e/ou gravação.
- .. discos giram em velocidades que variam de 3.600 rpm, p.ex., dispositivos portáteis, até 15.000 rpm e na qual uma revolução = 4 ms, ou seja, em média, o atraso rotacional de 2 ms.
- “**transfer time**” .. tempo de transferência de / para o disco depende da velocidade de rotação nominal do disco = $b / r * N$
- onde b = número de bytes a serem transferidos;
- N = número de bytes em uma trilha;
- r = velocidade de rotação, em rotações por segundo.

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.1 – Disk Performance Parameters

- “**disk input/output operation**” .. depende do sistema do computacional, do sistema operacional e da natureza do canal de I/O e do hardware do controlador de disco.
- “**average access time**” .. tempo total de acesso pode ser expresso ..
- $T_a = T_s + (1 / 2 * r) + (b / r * N)$
- onde T_s é o tempo médio de busca.



11.5.2 – Disk Scheduling Policies

- **“reason for the difference in performance”** .. está diretamente relacionada com o “seek time” – “tempo de busca”.
- .. se as solicitações de acesso ao setor envolve seleção aleatória de trilhas, o desempenho do sistema de I/O será o mais fraco possível.
- .. para melhorar, precisa-se reduzir o tempo médio gasto em buscas.
- e.g., considere a situação típica em um ambiente de multiprogramação, no qual o sistema operacional mantém uma fila de solicitações para cada dispositivo de I/O.
- .. portanto, para um único disco, haverá várias solicitações de I/O (leituras e gravações) de vários processos na fila.
- .. se selecionarmos itens da fila em ordem aleatória, espera-se que as trilhas sejam visitadas aleatoriamente » desempenho ruim.

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

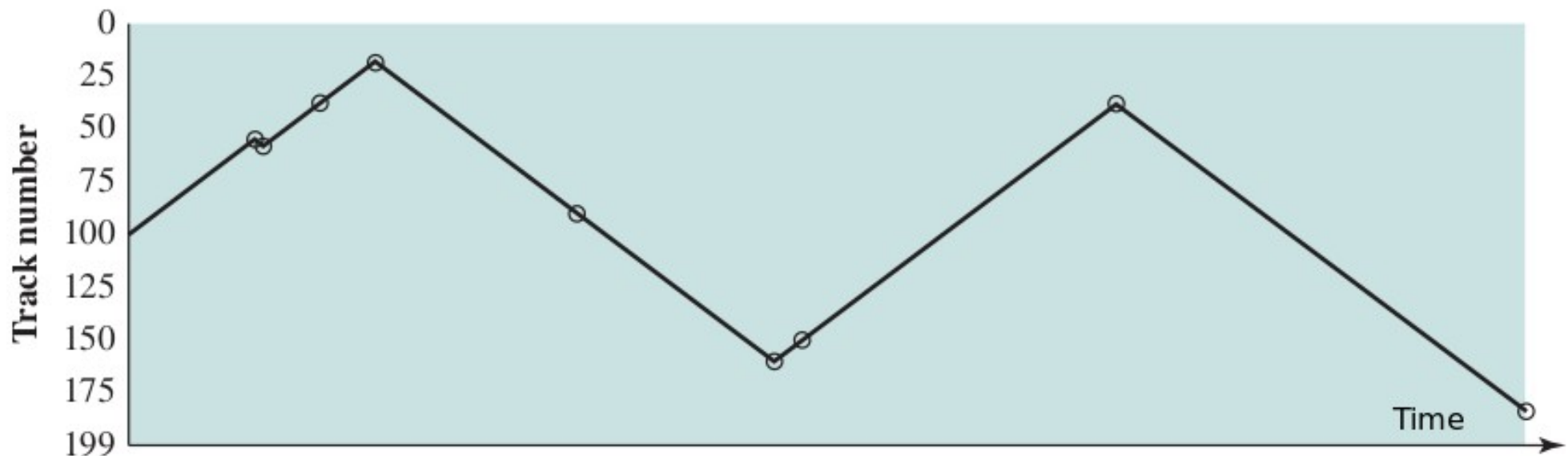
- e.g., tabela abaixo compara o desempenho de vários algoritmos de escalonamento para uma sequência de solicitações de I/O em um disco com 200 trilhas e cabeçote do disco localizado na trilha 100.
- .. supõe-se que a fila de solicitações do disco possui solicitações aleatórias, sendo que, as faixas solicitadas, na ordem recebida pelo escalonador de disco, são 55, 58, 39, 18, 90, 160, 150, 38, 184.

First In First Out (FIFO) (starting at track 100)		Shortest Service Time First (SSTF) (starting at track 100)		SCAN (starting at track 100, in the direction of increasing number)		C-SCAN (starting at track 100, in the direction of increasing number)	
Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed
55	45	90	10	150	50	150	50
58	3	58	32	160	10	160	10
39	19	55	3	184	24	184	24
18	21	39	16	90	94	18	166
90	72	38	1	58	32	38	20
160	70	18	20	55	3	39	1
150	10	150	132	39	16	55	16
38	112	160	10	38	1	58	3
184	146	184	24	18	20	90	32
Average Seek Length	55.3	Average Seek Length	27.5	Average Seek Length	27.8	Average Seek Length	35.8

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

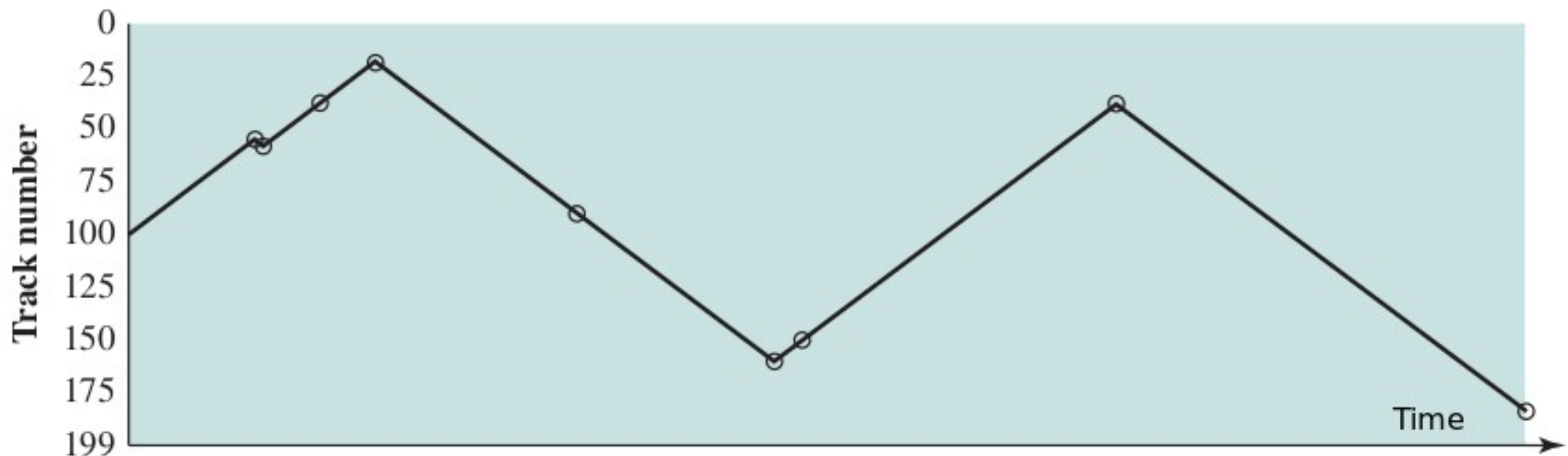
- **“First In First Out - FIFO”** .. forma mais simples de escalonamento e que processa os itens da fila em ordem sequencial.
- .. essa estratégia tem a vantagem de ser justa, pois todo pedido é atendido e os pedidos são atendidos na ordem recebida.
- .. como pode ser visto, os acessos ao disco estão na mesma ordem em que as solicitações foram recebidas originalmente.



11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

- .. se houver poucos processos competindo pelo disco e se muitas das solicitações forem para setores de arquivos em “cluster”, pode-se esperar um bom desempenho.
- .. se houver muitos processos competindo pelo disco, essa técnica aproxima-se do “random scheduling” em desempenho, assim, é mais vantajoso considerar uma política de escalonamento mais sofisticada



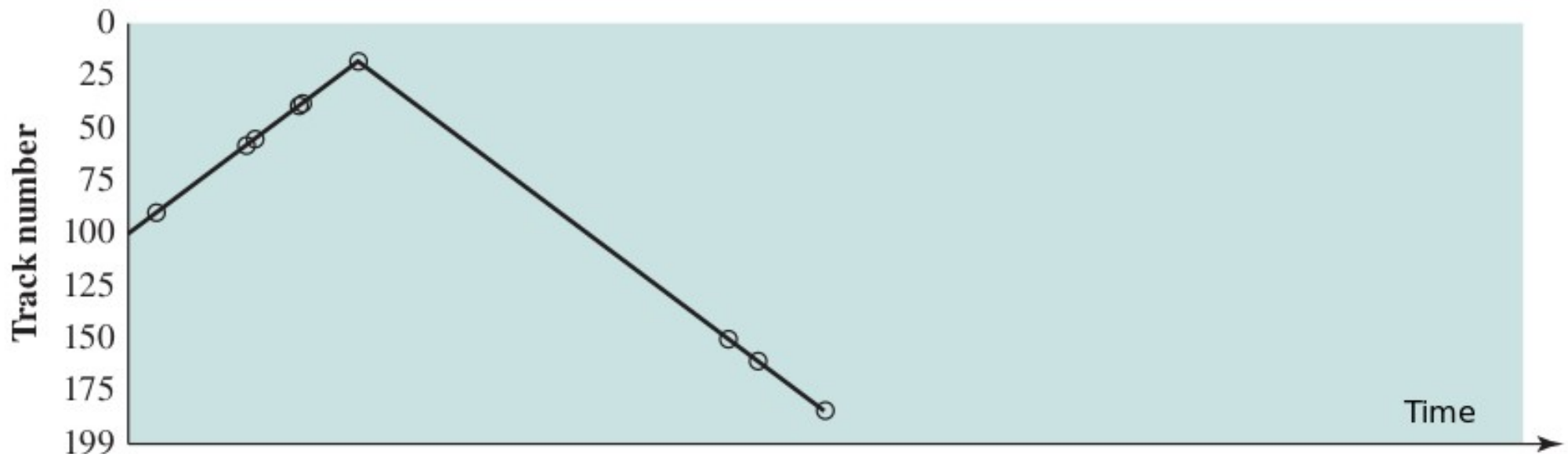
... 11.5.2 – Disk Scheduling Policies

- **“Priority”** .. com um sistema baseado em prioridade, o escalonamento está fora do controle do software de gerenciamento de disco.
- .. essa abordagem não visa otimizar a utilização do disco, mas atender a outros objetivos dentro do sistema operacional.
- e.g, muitas vezes, os trabalhos em lote curtos e os trabalhos interativos recebem prioridade mais alta do que os trabalhos mais longos que exigem computação mais longa.
- .. isso permite que muitos trabalhos curtos sejam liberados rapidamente pelo sistema e pode fornecer um bom tempo de resposta interativa, no entanto, trabalhos mais longos podem ter que esperar muito tempo.
- .. esta política pode levar a contramedidas por parte dos usuários, que dividem seus trabalhos em partes menores para superar o sistema.

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

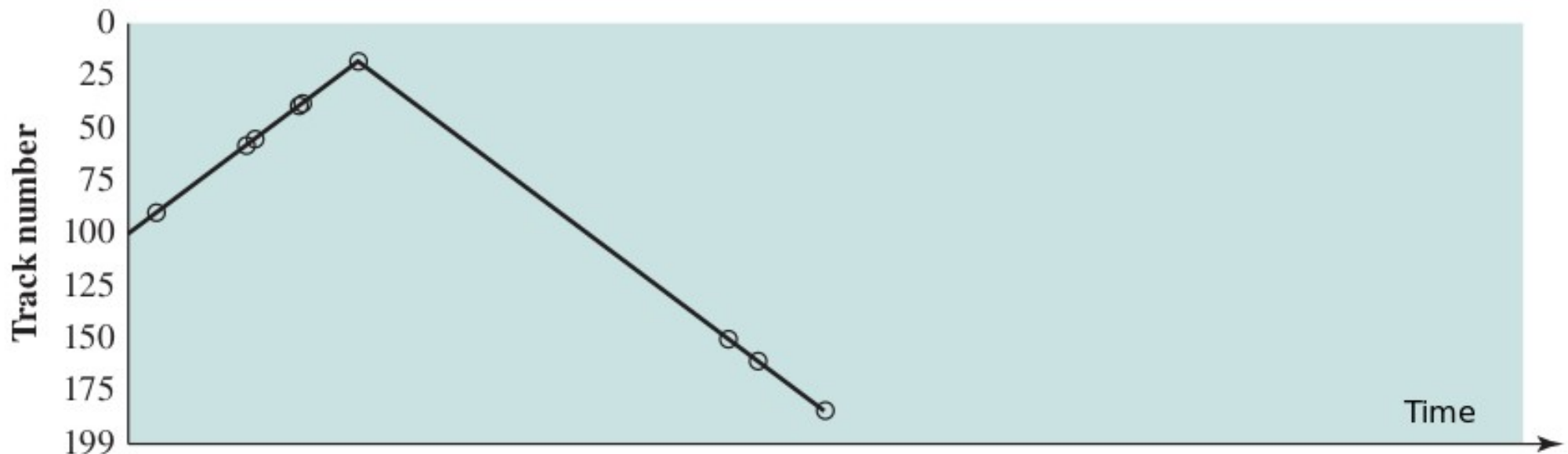
- **“Shortest Service Time First”** .. seleciona a solicitação de I/O que requer o menor movimento do braço de disco de sua posição atual, assim, escolhe-se o tempo mínimo de busca.
- .. obviamente, sempre escolher o tempo mínimo de busca não garante que o tempo médio de busca em vários movimentos seja mínimo, mas deve fornecer melhor desempenho do que FIFO.



11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

- A Figura (abaixo) e a Tabela (próximo slide) mostram o desempenho do SSTF no mesmo exemplo usado para FIFO.
- .. primeira trilha acessada é 90, pois esta é a trilha solicitada mais próxima da posição inicial e a próxima faixa acessada é 58 porque esta é a mais próxima das demais faixas solicitadas da posição atual de 90.



11 – Input / Output Management / 11.5 – Disk Scheduling

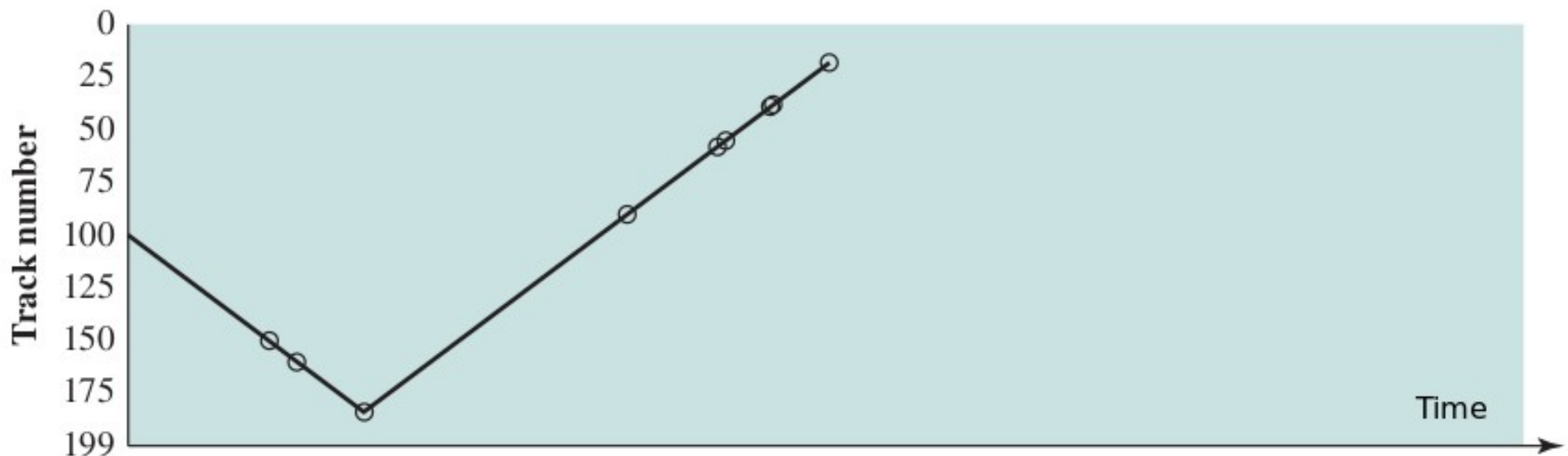
... 11.5.2 – Disk Scheduling Policies

- A Figura (slide anterior) e a Tabela mostram o desempenho do SSTF no mesmo exemplo usado para FIFO.
- .. primeira trilha acessada é 90, pois esta é a trilha solicitada mais próxima da posição inicial e a próxima faixa acessada é 58 porque esta é a mais próxima das demais faixas solicitadas da posição atual de 90.

First In First Out (FIFO) (starting at track 100)		Shortest Service Time First (SSTF) (starting at track 100)		SCAN (starting at track 100, in the direction of increasing number)		C-SCAN (starting at track 100, in the direction of increasing number)	
Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed
55	45	90	10	150	50	150	50
58	3	58	32	160	10	160	10
39	19	55	3	184	24	184	24
18	21	39	16	90	94	18	166
90	72	38	1	58	32	38	20
160	70	18	20	55	3	39	1
150	10	150	132	39	16	55	16
38	112	160	10	38	1	58	3
184	146	184	24	18	20	90	32
Average Seek Length	55.3	Average Seek Length	27.5	Average Seek Length	27.8	Average Seek Length	35.8

... 11.5.2 – Disk Scheduling Policies

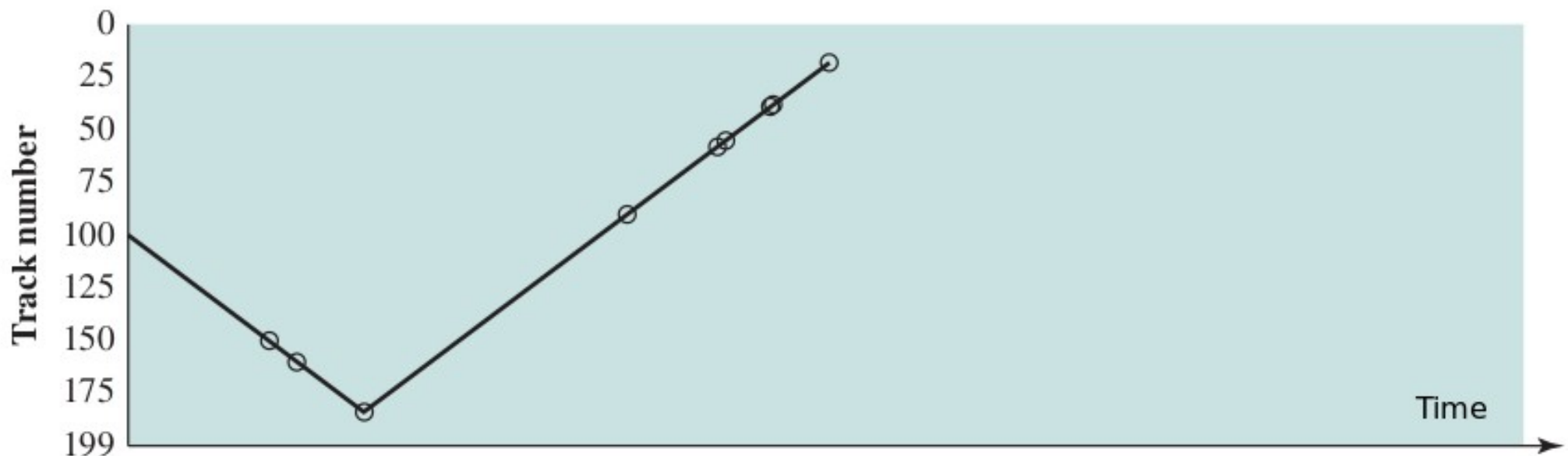
- “**SCAN**” .. braço move-se em apenas uma direção, satisfazendo todas as solicitações pendentes no caminho, até atingir a última faixa nessa direção ou até que não haja mais solicitações nessa direção.
- .. na sequência, a direção do serviço é então invertida e a varredura prossegue na direção oposta, atendendo as solicitações em ordem.



11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

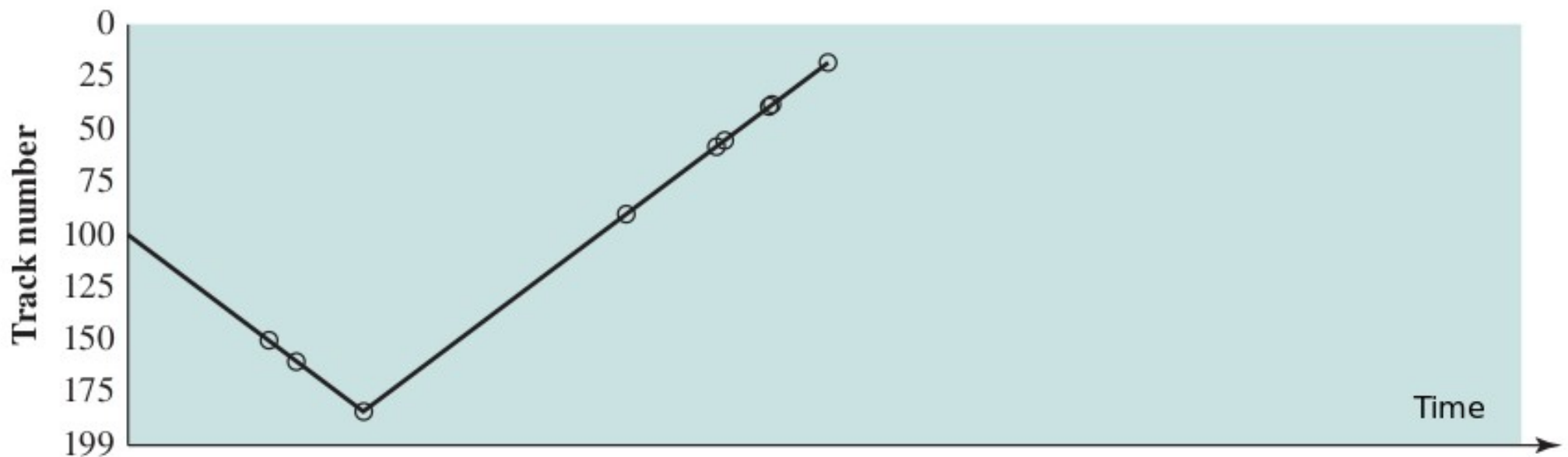
- SCAN comporta-se de forma quase idêntica ao SSTF e, caso o braço estivesse se movendo na direção dos nros. de faixa mais baixos, o padrão de escalonamento teria sido idêntico para SSTF e SCAN.
- .. mesmo quando a fila mudar dinamicamente, o SCAN será semelhante ao SSTF, a menos que o padrão de solicitação seja incomum.



11 – Input / Output Management / 11.5 – Disk Scheduling

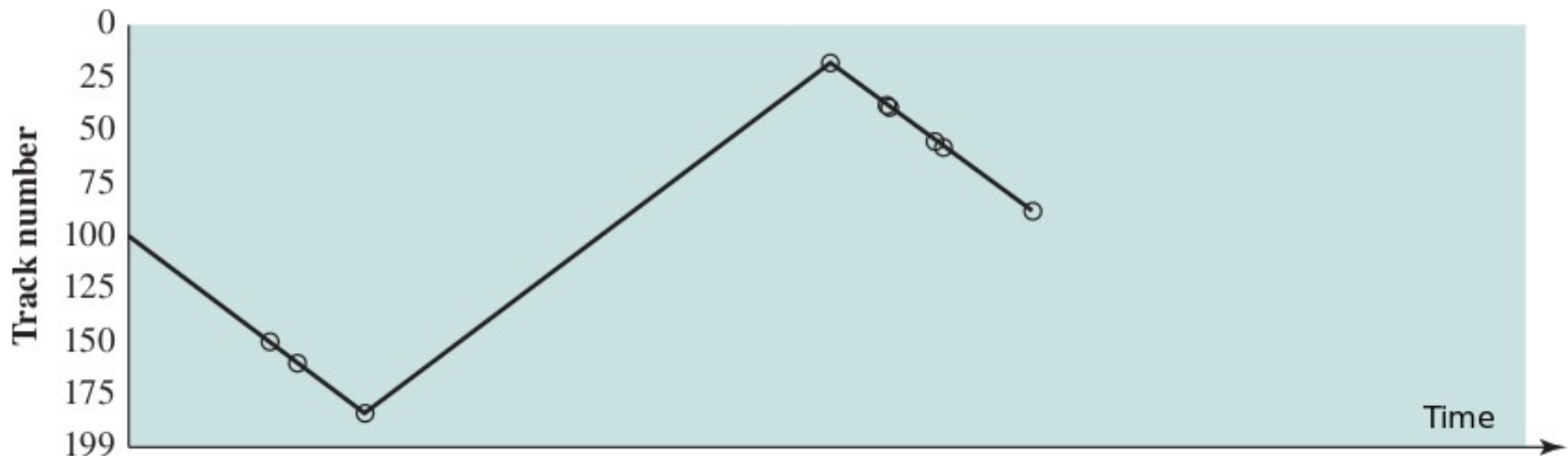
... 11.5.2 – Disk Scheduling Policies

- “**observação**” .. SCAN é tendenciosa contra a área mais recentemente atravessada, pois não explora a localidade tão bem quanto o SSTF.
- .. não é difícil perceber que o SCAN favorece os “jobs” cujas solicitações são para as trilhas mais próximas das trilhas mais internas e externas e favorece os trabalhos que chegam mais recentes.



... 11.5.2 – Disk Scheduling Policies

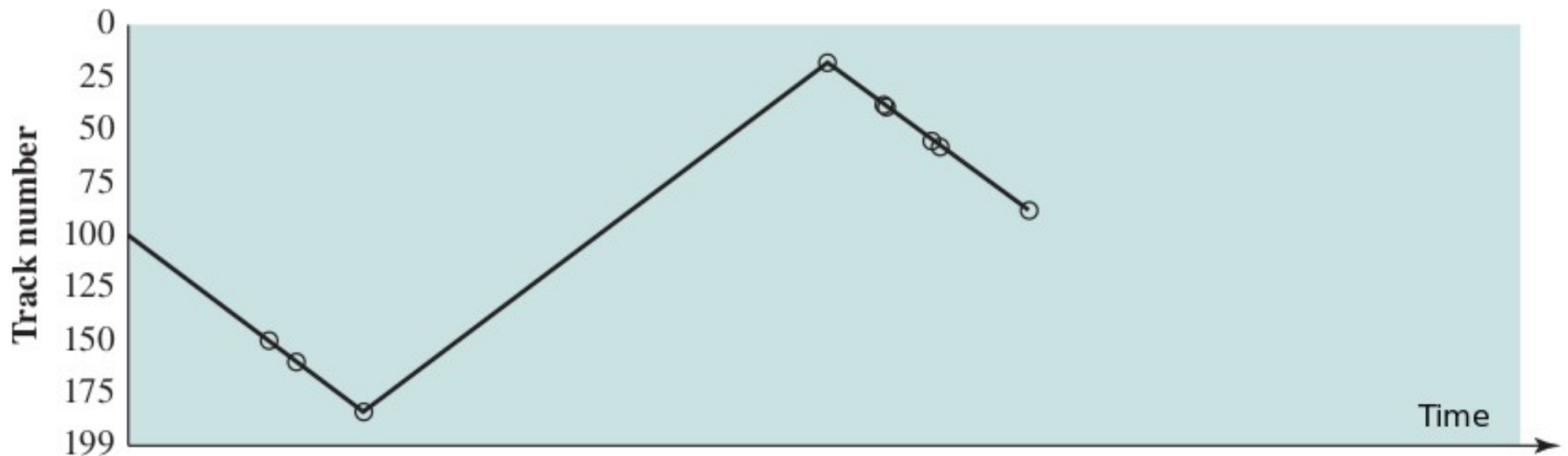
- **“C-SCAN”** .. restringe a varredura a apenas uma direção, assim, quando a última trilha é visitada em uma direção, o braço é devolvido à extremidade oposta do disco e a varredura começa novamente.
- .. nesse caso, as 03 primeiras faixas solicitadas encontradas são 150, 160 e 184. Em seguida, a varredura começa no número de faixa mais baixo e a próxima faixa solicitada encontrada é 18.



11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

- Com o SCAN, se o tempo esperado para uma varredura da trilha interna para a trilha externa for “ t ”, então o intervalo de serviço esperado para setores em a periferia é “ $2 * t$ ”.
- Com o C-SCAN, o intervalo é da ordem de “ $t + s_{max}$ ”, onde “ s_{max} ” é o tempo máximo de busca.



11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

- Com o SCAN, se o tempo esperado para uma varredura da trilha interna para a trilha externa for “t”, então o intervalo de serviço esperado para setores em a periferia é “2 * t”.
- Com o C-SCAN, o intervalo é da ordem de “t + s max” , onde “s max” é o tempo máximo de busca.

First In First Out (FIFO) (starting at track 100)		Shortest Service Time First (SSTF) (starting at track 100)		SCAN (starting at track 100, in the direction of increasing number)		C-SCAN (starting at track 100, in the direction of increasing number)	
Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed	Next Track Accessed	Number of Tracks Traversed
55	45	90	10	150	50	150	50
58	3	58	32	160	10	160	10
39	19	55	3	184	24	184	24
18	21	39	16	90	94	18	166
90	72	38	1	58	32	38	20
160	70	18	20	55	3	39	1
150	10	150	132	39	16	55	16
38	112	160	10	38	1	58	3
184	146	184	24	18	20	90	32
Average Seek Length	55.3	Average Seek Length	27.5	Average Seek Length	27.8	Average Seek Length	35.8

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

- SSTF, SCAN e C-SCAN .. se um ou alguns processos têm altas taxas de acesso a uma trilha, é possível a monopolização do dispositivo por meio de solicitações repetidas a essa trilha.
- .. discos multisuperfícies de alta densidade são mais propensos a serem afetados por essa característica do que discos de baixa densidade e/ou discos com apenas uma ou duas superfícies.
- .. para evitar essa “aderência de braço”, a fila de solicitações de disco pode ser segmentada, com um segmento de cada vez sendo completamente processado » “N-step-SCAN” e “FSCAN”.

Selection According
to Requestor

Name	Description	Remarks
RSS	Random scheduling	For analysis and simulation
FIFO	First-in-first-out	Fairest of them all
PRI	Priority by process	Control outside of disk queue management
LIFO	Last in first out	Maximize locality and resource utilization

... 11.5.2 – Disk Scheduling Policies

- **“N-step-SCAN”** .. segmenta a fila de solicitações de disco em subfilas de comprimento N e, na sequência, as subfilas são processadas uma de cada vez, usando SCAN.
- .. enquanto uma fila está sendo processada, novas solicitações devem ser adicionadas a alguma outra fila.
- .. se menos de N solicitações estiverem disponíveis no final de uma varredura, todas elas serão processadas com o próxima varredura.
- .. com grandes valores de N, o “N-step-SCAN” se aproxima do SCAN e com um valor de $N = 1$ a política FIFO é adotada.

11 – Input / Output Management / 11.5 – Disk Scheduling

... 11.5.2 – Disk Scheduling Policies

- “**FSCAN**” .. política que usa 02 subfilas, assim, quando uma varredura começa, todas as solicitações estão em 01 das filas, com a outra vazia.
- .. durante a varredura, todas as novas solicitações são colocadas na outra fila, portergando o atendimento de novas requisições até que todas as requisições antigas tenham sido processadas.

Selection According to Requested Item	Name	Description	Remarks
	SSTF	Shortest-service-time first	High utilization, small queues
	SCAN	Back and forth over disk	Better service distribution
	C-SCAN	One way with fast return	Lower service variability
	<i>N</i> -step-SCAN	SCAN of <i>N</i> records at a time	Service guarantee
	FSCAN	<i>N</i> -step-SCAN with <i>N</i> = queue size at beginning of SCAN cycle	Load sensitive

11 – Input / Output Management / 11.6 – RAID

11.6 – Redundant Array of Independent Disks

- aaa

11 – Input / Output Management / 11.7 – Disk Cache

11.7 – Disk Cache

- aaa

11 – Input / Output Management / 11.7 – Disk Cache

11.7.1 – Design Considerations

- aaa

11 – Input / Output Management / 11.7 – Disk Cache

11.7.2 – Performance Considerations

- aaa

11 – Input / Output Management / 11.8 – UNIX SVR4 Input/Output

11.8 – UNIX SVR4 Input/Output

- .. LEITURA COMPLEMENTAR !!!

11 – Input / Output Management / 11.9 – LINUX Input/Output

11.9 – LINUX Input/Output

- .. LEITURA COMPLEMENTAR !!!

11 – Input / Output Management / 11.10 – WINDOWS Input/Output

11.10 – WINDOWS Input/Output

- .. LEITURA COMPLEMENTAR !!!