< VOLTAR



Escalonamento de CPU

Neste tópico, iremos estudar as formas como o sistema operacional escolhe (ou é configurado para executar) os processos para serem executados.

NESTE TÓPICO

- > Introdução
- > Escalonamento
- > Referências

Marcar tópico





Introdução

Atualmente os SO são multitarefa / multiprogramação, porém, apenas uma instrução de processo é executado por vez nos computadores, ou seja, o processador divide "sua atenção" a espaços de tempo dedicados a cada processo, atendendo todos, porém, uma instrução de cada um por vez.

Escalonamento

Desde a criação dos computadores, houve a preocupação com a otimização de recursos. Executar as tarefas nos computadores sempre no menor tempo possível, foi sempre uma prioridade. Segundo Machado (2007), a política de escalonamento de um sistema operacional tem diversas funções básicas:

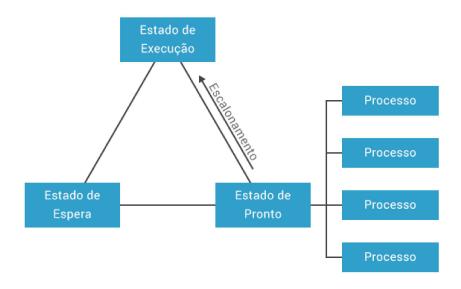
- Manter o processador ocupado a maior parte do tempo
- Balancear o uso de CPU entre processos;
- Privilegiar a execução de aplicações críticas;
- Maximizar o throughput do Sistema;
- Oferecer tempo de respostas variáveis para os usuários interativos.

A principal função do SO é implementar os critérios da política de escalonamento, denominada escalonador (scheduler). Escalonamento é a priorização de processamentos.

Em um sistema multiprogramável, o escalonador é fundamental, porque sem ele não é possível o compartilhamento do processador, pois é por meio desse algoritmo que o sistema operacional consegue gerenciar os processos que fazem uso do processador.

Há duas rotinas importantes para a gerência do processador, são elas:

- **Dispatcher**: responsável pela troca de contexto dos processos após o escalonador determinar qual processo deve fazer uso do processador;
- Latência de Dispatcher: tempo gasto na substituição de um processo em execução por outro.



Escalonamento.

Fonte: MACHADO, F. B.; MAIA, L. P. Arquitetura de Sistemas Operacionais. 2. ed. LTC, 2002.

Critérios de Escalonamento

De acordo com as características do sistema operacional, determinam-se quais são os atributos para a implementação de um escalonamento adequado. Segundo Machado (2007), os critérios são:

- Utilização do processador: é desejável que o processador permaneça a maior parte do seu tempo ocupado. Uma utilização na faixa de 30% indica um sistema com uma carga de processamento baixa, enquanto na faixa de 90% indica um sistema bastante carregado, próximo da capacidade máxima;
- Throughput: representa o número de processos executados em um determinado intervalo de tempo. Quanto maior o throughput, maior o

número de tarefas executadas em função do tempo. A maximização do throughput é desejada na maioria dos sistemas operacionais;

- Tempo de processador / tempo de CPU: é o tempo que um processo leva no
 estado de execução durante seu processamento. As políticas
 de escalonamento não influenciam o tempo de processador de um
 processo, sendo este tempo função apenas do código da aplicação e da
 entrada de dados;
- Tempo de espera: é o tempo total que um processo permanece na fila de pronto durante seu processamento aguardando para ser executado. A redução do tempo de espera dos processos é desejada pela maioria das políticas de escalonamento;
- Tempo de turnaround: é o tempo que um processo leva desde a sua criação até seu término, levando em consideração todo o tempo gasto na espera para alocação de memória, espera na fila de pronto, processamento de CPU e na fila de espera, como nas operações de E/S. as políticas de escalonamento buscam minimizar o tempo de turnaround;
- Tempo de resposta: é o tempo decorrido entre uma requisição ao sistema ou à aplicação e o instante em que a resposta é exibida. Em sistemas interativos, pode-se entender como o tempo decorrido entre a última tecla digitada pelo usuário e o início da execução do resultado no monitor.

Escalonamentos não preemptivos e preemptivos

Podemos classificar os escalonamentos de acordo com a possibilidade ou não da interrupção de um processo.

Se o sistema operacional consegue interromper um processo em execução e substituí-lo por outro, trata-se de um escalonamento do **Tipo Preempção**. Esses sistemas são mais complexos, porém possibilitam políticas de escalonamentos mais flexíveis.

Há também o E**scalonamento não Preemptivo**, o primeiro implementado nos sistemas operacionais multiprogramáveis. Nesse modelo, quando um processo está em execução, nenhum evento externo irá ocasionar a perda do uso do processador, isso apenas acontecerá se o processo concluir seu processamento ou executar instruções do próprio código que promovam uma mudança de estado de espera.

A vantagem do escalonamento preemptivo é que o sistema operacional pode interromper um processo em execução e passá-lo para o estado de pronto, no intuito de alocar outro processo na CPU. Com a preempção, é possível ao sistema priorizar a execução de processos, como nas aplicações de tempo real, onde o fator tempo é crítico. Outro benefício é a possibilidade de implementar políticas de escalonamento que compartilhem o processador de maneira mais uniforme, distribuindo de forma balanceada o uso da CPU entre os processos, aumentando a produtividade e diminuindo o tempo de turnaround.

Escalonamento First-Come-First-Served (FCFS)

Conhecido também como FIFO (First-In_First_Out: o primeiro a entrar é o primeiro a sair). O conceito desse algoritmo de escalonamento, o processo que chegar primeiro ao estado de pronto é escolhido para execução.

É necessária apenas uma fila (atendidos como uma fila no caixa de um banco), onde os processos que passam para o estado de pronto entram no final e são escolhidos quando chegam ao seu início. Se um processo vai para o estado de espera, o primeiro que estiver na fila de pronto é escalonado.

O grande problema é prever quando um processo terá sua execução iniciada, pois isso depende dos que estão na frente da fila.

Esse escalonamento é do tipo não preemptivo, ou seja, não interrompe, e foi inicialmente desenvolvido e implementado em sistemas monoprogramáveis com processamento batch, sendo inviável aplicar esse conceito em sistemas de compartilhado.

Escalonamento Shortest Job First (SJF)

Esse escalonamento atende primeiramente os processos menores, aqueles que menos utilizam a CPU. O cálculo de cada tempo médio é feito a partir de uma segunda alocação de CPU, ou seja, o processo que utilizar a CPU por menos tempo será executado primeiro. Foi desenvolvido inicialmente para os sistemas operacionais batch (processamento em lotes). Esse tipo de escalonamento pode ser preemptivo e não preemptivo.

Escalonamento SJF não Preemptivo

Nesse escalonamento, o processo faz uso do processador até o final da sua execução sem interrupção. Nenhum evento externo influenciará essa execução. Ao concluir o processamento, verifica-se na fila de pronto qual é o próximo "menor processo" a ser executado.

Escalonamento SJF Preemptivo

Nesse modelo, caso chegue um processo menor que o processo que está sendo executado, o "maior" (que utiliza mais CPU) será interrompido, para atender o "menor" que chegou.

Chegamos ao final deste tópico, esperamos que você tenha gostado e absorvido o conhecimento dos assuntos abordados..

Compreender a como o computador executa os aplicativos (Processos: programas em execução) é muito importante! Esperamos que você tenha entendido bem o que é isso. Caso haja alguma dúvida, não deixe de leva-la ao professor, ok?

Faça os exercícios e verifique seu conhecimento.

Bom estudo!



Escalonamento de CPU

INICIAR >

Quiz

Exercício Final

Escalonamento de CPU

INICIAR >

Referências

DEITEL, H. M.; DEITEL, P. J.; CHOFFNES, D. R. Sistemas operacionais. 3. ed. São Paulo: Pearson Prentice Hall, 2005.

MACHADO, F. B.; MAIA, L. P. Arquitetura de sistemas operacionais. 2. ed. São Paulo: LTC, 2002.

TANENBAUM, A. S. Sistemas operacionais modernos. 2. ed. São Paulo: Pearson Prentice Hall, 2005.



Avalie este tópico



ANTERIOR Threads

=

Biblioteca

Índice

(https://www.uninove.br/conheca-

а

uninove/biblioteca/sobre-

a-

biblioteca/apresentacao/)

Portal Uninove

(http://www.uninove.br)

Mapa do Site



® Todos os direitos reservados

