Aula 05 Sistemas Operacionais I

Processos - Parte 02

Prof. Julio Cezar Estrella

jcezar@icmc.usp.br

Material adaptado de

Sarita Mazzini Bruschi

baseados no livro Sistemas Operacionais Modernos de A. Tanenbaum

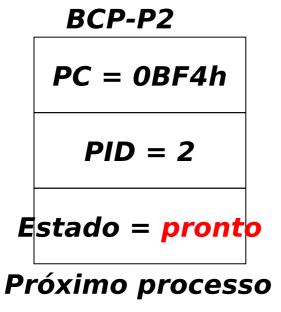
Processos

- Introdução
- Escalonamento de Processos
- Comunicação entre Processos
- Threads
- Deadlock

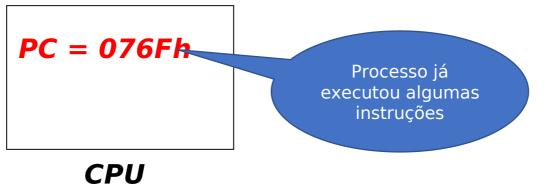
- Escalonador de Processos escolhe o processo que será executado pela CPU;
- Escalonamento é realizado com o auxílio do hardware;
- Escalonador deve se preocupar com a eficiência da CPU, pois o chaveamento de processos é complexo e custoso:
 - Afeta desempenho do sistema e satisfação do usuário;
- Escalonador de processo é um processo que deve ser executado quando da **mudança de contexto** (troca de processo);

- Mudança de Contexto:
 - Overhead de tempo;
 - Tarefa cara:
 - Salvar as informações do processo que está deixando a CPU em seu
 BCP → conteúdo dos registradores;
 - Carregar as informações do processo que será colocado na CPU → copiar do BCP o conteúdo dos registradores;

Antes da Mudança de Contexto







Depois da Mudança de Contexto



- Situações nas quais escalonamento é necessário:
 - Um novo processo é criado;
 - Um processo terminou sua execução e um processo pronto deve ser executado;
 - Quando um processo é bloqueado (semáforo, dependência de E/S), outro deve ser executado;
 - Quando uma interrupção de E/S ocorre o escalonador deve decidir por: executar o processo que estava esperando esse evento; continuar executando o processo que já estava sendo executado ou executar um terceiro processo que esteja pronto para ser executado;

- Hardware de relógio fornece interrupções de relógio e a decisão do escalonamento pode ser tomada a cada interrupção ou a cada k interrupções;
- Algoritmos de escalonamento podem ser divididos em duas categorias dependendo de como essas interrupções são tratadas:
 - Preemptivo: escolhe um processo e o deixa executando por um tempo máximo;
 - Não-preemptivo: estratégia de permitir que o processo que está sendo executado continue sendo executado até ser bloqueado por alguma razão (semáforos, operações de E/S-interrupção) ou que libere a CPU voluntariamente;

- Categorias de Ambientes:
 - Sistemas em Batch: usuários não esperam por respostas rápidas; algoritmos não-preemptivos ou preemptivos com longo intervalo de tempo;
 - Sistemas Interativos: interação constante do usuário; algoritmos preemptivos; Processo interativo → espera comando e executa comando;
 - Sistemas em Tempo Real: processos são executados para não deixar de cumprir o prazo que o processo tem que ser executado; tempo é crucial → sistemas críticos;

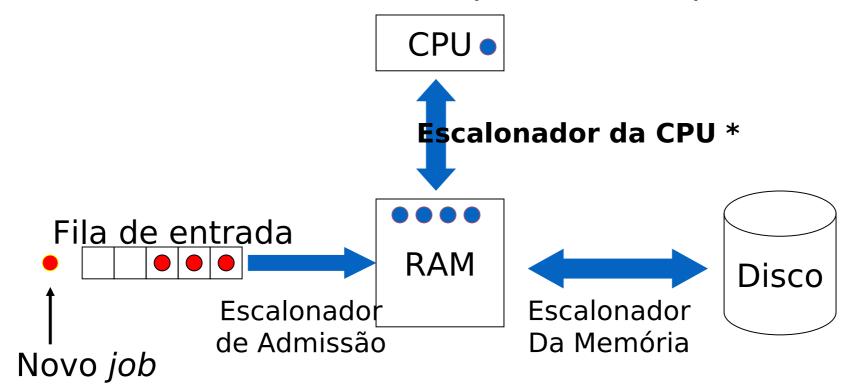
- Características de algoritmos de escalonamento:
 - Qualquer sistema:
 - <u>Justiça</u> (*Fairness*): cada processo deve receber uma parcela justa de tempo da CPU;
 - **Balanceamento**: diminuir a ociosidade do sistema;
 - Políticas do sistema: prioridade de processos;

- Características dos algoritmos de escalonamento:
 - Sistemas em *Batch*:
 - **Vazão** (throughput): maximizar o número de jobs executados por hora;
 - <u>Tempo de retorno</u> (*turnaround time*): tempo no qual o processo espera para ser finalizado;
 - **Eficiência**: CPU deve estar 100% do tempo ocupada;
 - Sistemas Interativos:
 - **Tempo de resposta**: tempo esperando para iniciar execução;
 - **Proporcionalidade**: satisfação do usuários;

- Características dos algoritmos de escalonamento:
 - Sistemas em Tempo Real:
 - **Cumprimento dos prazos**: prevenir perda de dados;
 - Previsibilidade: prevenir perda da qualidade dos serviços oferecidos;

- Algoritmos para Sistemas em Batch:
 - Três níveis
 - First-Come First-Served (ou FIFO)
 - Shortest Job First (SJF)
 - Shortest Remaining Time Next (SRTN)

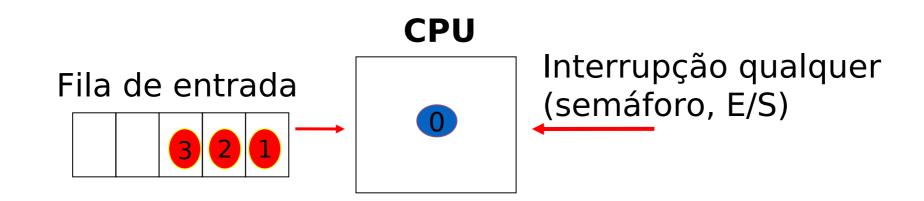
• Escalonamento *Três níveis* (*Three-level*)



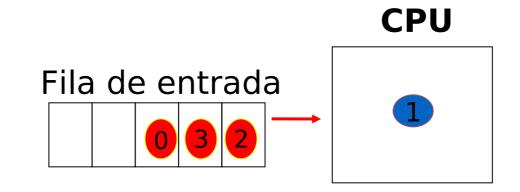
- Escalonamento *Três níveis*
 - Escalonador de admissão: decide qual job será admitido no sistema. Por exemplo, uma mescla de jobs orientados a CPU e orientados à E/S; processos com menor tempo de acesso à CPU e maior tempo de interação com dispositivos de E/S;
 - Escalonador da Memória: decisões sobre quais processos vão para a MP:
 - A quanto tempo o processo está esperando?
 - Quanto tempo da CPU o processo já utilizou?
 - Qual o tamanho do processo?
 - Qual a importância do processo?
 - Escalonador da CPU: seleciona qual o próximo processo a ser executado;

- Algoritmo First-Come, First-Served
 - Não-preemptivo;
 - Processos são executados na CPU seguindo a ordem de requisição;
 - Fácil de entender e programar;
 - <u>Desvantagem</u>:
 - Ineficiente quando se tem processos que demoram na sua execução;

Algoritmo First-Come, First-Served



Algoritmo First-Come, First-Served



CPU não controla o tempo dos processos! (não-preemptivo)

Algoritmo Shortest Job First

- Não-preemptivo;
- Possível prever o tempo de execução do processo;
- Menor processo é executado primeiro;
- Menor turnaround médio;
- <u>Desvantagem</u>:
 - Baixo aproveitamento quando se tem poucos processos prontos para serem executados;

Algoritmo Shortest Job First

```
A \rightarrow a
B \rightarrow b+a
C \rightarrow c+b+a
D \rightarrow d+c+b+a
```

Tempo médio-turnaround (4a+3b+2c+d)/4

Contribuição → se a<b<c<d tem-se o mínimo tempo médio;

Algoritmo Shortest Job First

8	4	4	4
Α	В	С	D

4	4	4	8
В	С	D	Α

Em ordem:

Turnaround A = 8

Turnaround B = 12

Turnaround C = 16

Turnaround D = 20

Média \rightarrow 56/4 = 14

Menor *job* primeiro:

Turnaround B = 4

Turnaround C = 8

Turnaround D = 12

Turnaround A = 20

Média \rightarrow 44/4 = 11

Algoritmo Shortest Remaining Time Next

- Preemptivo;
- Processos com menor tempo de execução são executados primeiro;
- Se um processo novo chega e seu tempo de execução é menor do que do processo corrente na CPU, a CPU suspende o processo corrente e executa o processo que acabou de chegar;
- <u>Desvantagem</u>: processos que consomem mais tempo podem demorar muito para serem finalizados se muitos processos pequenos chegarem!