2.5 - Escalonamento de Processos

- ... quando mais de um processo está ativo, cabe ao sistema operacional decidir qual terá a posse da CPU, ou seja, é o escalonador que toma esta decisão.

- ... vários critérios devem ser observados por um algoritmo de escalonamento:

  ❶ progresso - garantir que cada processo tenha acesso à CPU;
  ❷ eficiência - manter a CPU ocupada praticamente 100% do tempo;
  ❸ tempo de resposta - minimizar o tempo de resposta na execução dos processos, principalmente os interativos (editores, planilhas, etc.);
  ❹ tempo de espera - minimizar o tempo de espera de serviços não interativos (compilação, impressão, etc.);
  ❺ vazão - maximizar o número de processos executados por troca.

✍ ... cabe observar que alguns desses objetivos são contraditórios.

... 2.5 - Escalonamento de Processos

- **escalonamento round robin** - é caracterizado por permitir que cada processo execute por um intervalo de tempo, comumente denominado *quantum*;

- ... se o processo estiver ainda executando ao final do *quantum*, ele é suspenso e a CPU é alocada a outro processo;

- ... se o processo for bloqueado antes do final do *quantum*, também neste caso a CPU é passada a outro processo;

✭ ... qual deve ser o tamanho do *quantum* ?!

  - se muito pequeno, diminui a eficiência da CPU pois causa *overhead*;
  - se for muito grande, a resposta para processos interativos é degradada.

... 2.5 - Escalonamento de Processos

- **algoritmos com prioridade** - diferentemente do **round robin**, estes algoritmos possibilitam a atribuição de prioridades aos processos;

- ... a idéia é que cada processo tem uma prioridade e processos com prioridades superiores devem ser executadas primeiro.

✯ como então impedir que processos com alta prioridade executem indefinidamente ?!

- ... via de regra, o escalonador diminui a prioridade dos processos com o aumento de seu respectivo tempo de execução.

... 2.5 - Escalonamento de Processos

- **múltiplas filas** - semelhante ao algoritmo anterior, este algoritmo define no entanto classes de prioridades para os processos;

- ... processos pertencentes à menor classe de prioridades executa por um $quantum$, processos da classe seguinte por 2 $quanta$, processos da próxima classe por 4 $quanta$ e assim por diante;

- ... quando um processo utiliza todos os $quanta$ a ele alocados, o mesmo é interrompido e sua classe tem a prioridade diminuída;

✔ este algoritmo diminui o número de comutações da CPU entre os processos ativos.

... 2.5 - Escalonamento de Processos

- **tarefas pequenas primeiro** - designado para aplicações interativas, este algoritmo necessita do conhecimento *a priori* do tempo médio de execução dos processos;
- ... neste algoritmo, tarefas menores são executada primeiro.
- **algoritmo policy-driven** - a CPU é particionada de forma equânime entre os usuários e não entre os processos, ou seja, em um sistema com n usuários, cada usuário deverá receber $1/n$ do poder da CPU;
- ... por outro lado, o sistema precisa manter informações do tempo de CPU que cada usuário já dispôs desde que entrou no sistema, bem como do instante de tempo que cada usuário ligou-se ao sistema.

... 2.5 - Escalonamento de Processos

- **escalonamento em dois níveis** - considera que nem todos os processos ativos residem na memória primária e, assim, define níveis de escalonamento associados à memória primária e secundária;
- ... um escalonador de baixo nível se restringe à troca de processos que estão na memória primária no momento;
- ... um escalonador de alto nível decide sobre a troca dos processos entre as memórias primária e secundária.