



Estudo Dirigido:

Escalonamento de CPU

Aluno: Victor Souza Santos
R.A.: 00000844512
Professor: Marcos Canejo
Curso: Ciências da Computação

1. Conceitos Básicos

- **O que é Escalonamento de CPU**

Escalonamento de CPU é o processo pelo qual um sistema operacional decide qual dos processos prontos na fila de processos deve ser executado pela CPU. Este mecanismo é essencial em sistemas operacionais multitarefa para garantir que todos os processos recebam tempo adequado de CPU para execução. O escalonamento é necessário para:

- Maximizar a utilização da CPU.
- Melhorar a eficiência do sistema.
- Garantir a equidade na distribuição dos recursos de CPU entre processos.
- Minimizar tempos de espera, resposta e turnaround (tempo de conclusão).

- **Ciclo de Picos de CPU-I/O**

Os processos alternam entre estados de execução de CPU e espera por I/O. Esse ciclo é ilustrado na Figura 1.

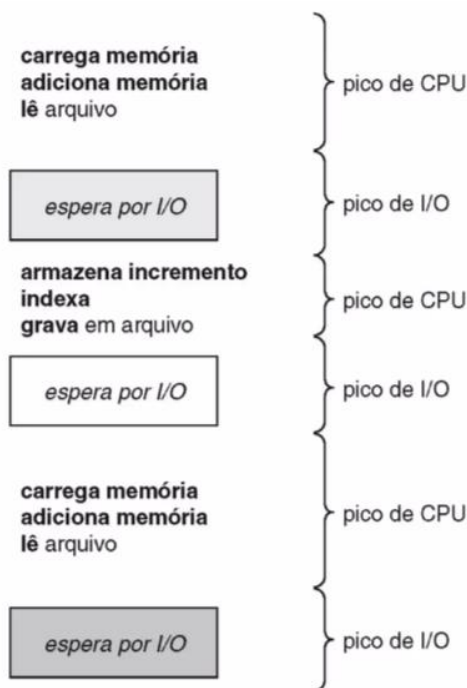


Figura 1: A execução de um processo é um ciclo de rajadas alternadas de CPU e espera por I/O

- **Scheduling com e sem Preempção**

- *Sem Preempção*: Uma vez que um processo começa a executar, ele continua até completar sua execução ou entrar em estado de espera (por exemplo, para I/O). Exemplos incluem FCFS e SJF não preemptivo.

- *Com Preempção*: A CPU pode ser retirada de um processo em execução para permitir que outro processo de maior prioridade ou que tenha chegado mais recentemente seja executado. Exemplos incluem Round-Robin e SJF preemptivo.

- **Despachante**

O despachante é um módulo do sistema operacional responsável por ceder o controle da CPU ao processo selecionado pelo escalonador. Suas funções incluem:

- Troca de contexto.
- Alternância para modo usuário.
- Retomada do ponto de execução correto no programa do usuário.

2. Critérios de Escalonamento

- **A. Utilização da CPU**

Este critério mede a proporção do tempo em que a CPU está ativa, ou seja, executando processos úteis. Idealmente, a CPU deve estar ocupada quase todo o tempo. A Figura 2 ilustra o conceito.

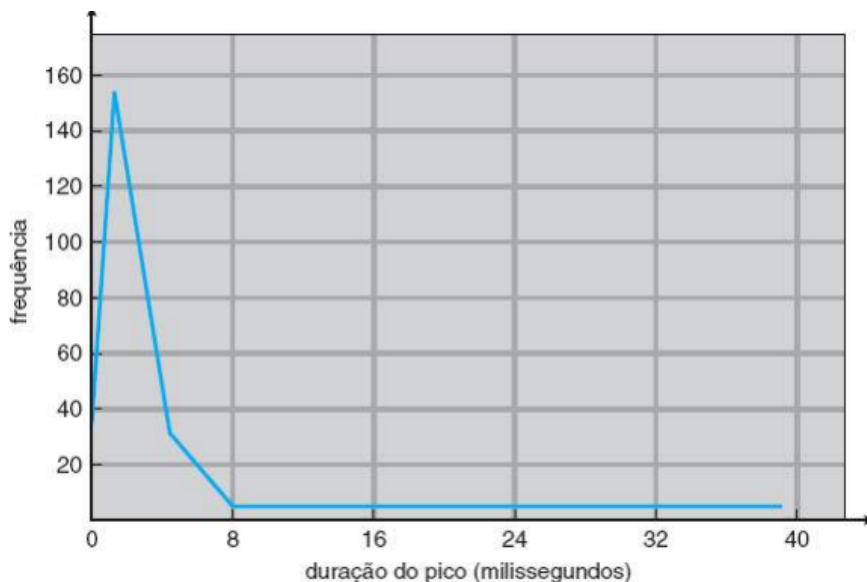


Figura 2: Histograma de

durações de picos de CPU.

- **B. Taxa de Transferência**

Refere-se ao número de processos concluídos por unidade de tempo. Algoritmos eficientes devem maximizar essa taxa.

- **C. Tempo de Espera**

O tempo de espera é o total de tempo que um processo passa na fila de prontos antes de ser executado. Minimizar este tempo melhora a performance percebida pelos usuários.

- **D. Tempo de Resposta**

É o tempo que se passa desde a submissão de um processo até a primeira resposta produzida. Tem um impacto direto em sistemas interativos.

- **E. Tempo de Turnaround**

Refere-se ao tempo total para completar um processo, desde a submissão até a conclusão. É a soma do tempo de execução e do tempo de espera.

3. Scheduling para Múltiplos Processadores

- **Funcionamento Geral**

Em sistemas com múltiplos processadores, o escalonamento envolve distribuir a carga de trabalho entre as CPUs disponíveis. Isso pode ser feito de duas maneiras principais:

- *Escalonamento Simétrico Multiprocessador (SMP)*: Todos os processadores compartilham a fila de prontos e executam o escalonamento de forma cooperativa.

- *Escalonamento Assimétrico Multiprocessador (AMP)*: Um processador mestre controla o escalonamento e distribui as tarefas para os processadores escravos. A Figura 3 mostra uma configuração SMP.

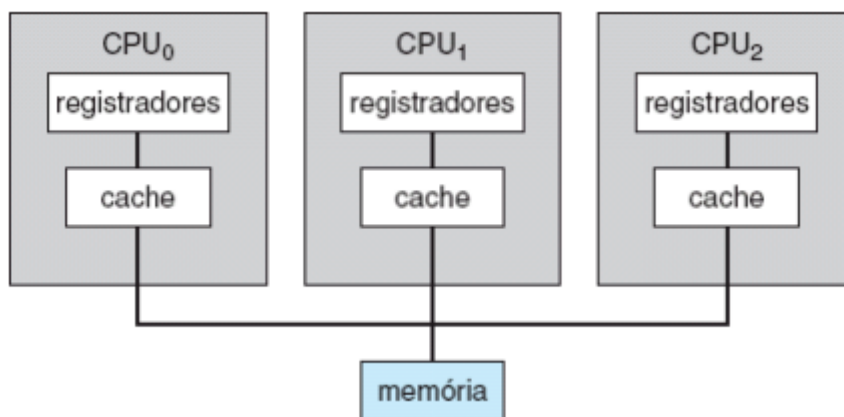


Figura 3: Arquitetura de

multiprocessamento simétrico.

- **Afinidade com o Processador**

A afinidade com o processador refere-se à preferência de um processo em ser executado pelo mesmo processador que o executou anteriormente. Isso pode melhorar a eficiência devido ao cache de dados local. A afinidade pode ser de dois tipos:

- *Afinidade Rígida*: O processo é fixado em um único processador.
- *Afinidade Suave*: O processo tem preferência, mas pode ser movido se necessário.

- **Balanceamento de Carga**

O balanceamento de carga garante que todos os processadores tenham uma carga de trabalho aproximadamente igual. Técnicas comuns incluem:

- *Balanceamento Centralizado*: Um único processador ou controlador redistribui a carga.

- *Balanceamento Distribuído*: Cada processador ou um grupo de processadores monitoram a carga e redistribuem tarefas conforme necessário.

A Figura 4 ilustra o balanceamento de carga.

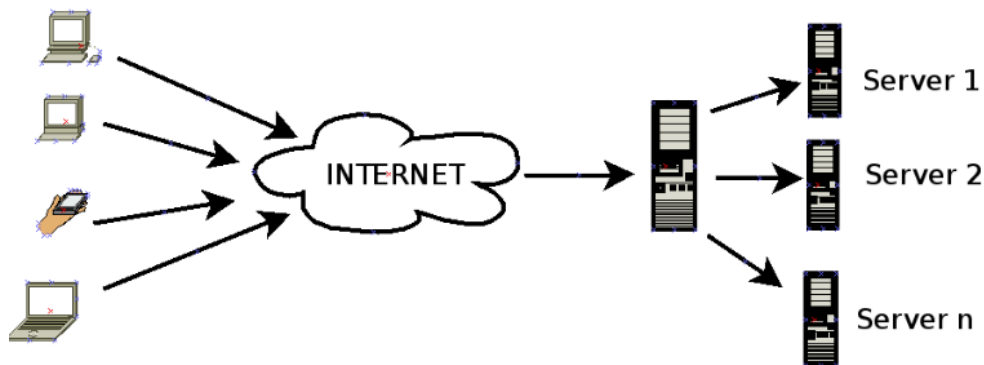


Figura 4: Exemplo de

balanceamento de carga centralizado.

4. Conclusão

O escalonamento de CPU é uma área fundamental para a eficiência dos sistemas operacionais multitarefa. Diferentes algoritmos oferecem vantagens específicas e são adequados para diferentes tipos de sistemas e cargas de trabalho. A escolha do algoritmo de escalonamento deve considerar critérios como utilização da CPU, taxa de transferência, tempo de espera, tempo de resposta e tempo de turnaround. Em sistemas multiprocessadores, a afinidade com o processador e o balanceamento de carga são cruciais para manter uma operação eficiente.

5. Referências

- Silberschatz, A., Galvin, P. B., & Gagne, G. (2018). Operating System Concepts. John Wiley & Sons.
- Stallings, W. (2014). Operating Systems: Internals and Design Principles. Pearson Education.