

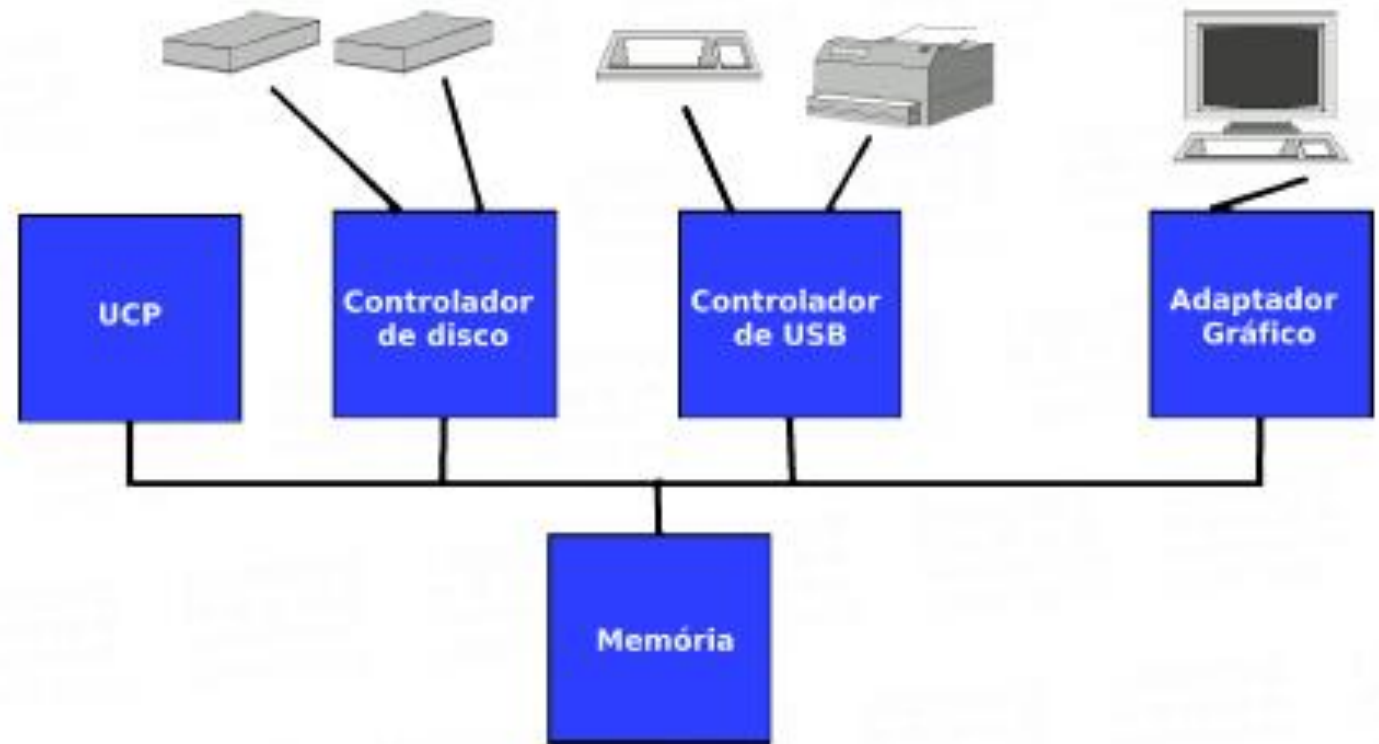


Processos - Sistemas Operacionais

Prof. Me. Hélio Esperidião

Organização
do
computador

Organização do Computador



Inicialização

Bootstrap ou Programa de Inicialização

Armazenado em firmware (ROM, EEPROM)

Inicializa e testa hardware

Carrega o núcleo do sistema operacional (kernel)

Primeiros processos são carregados

Espera-se a ocorrência de algum evento

Ocorrência de Eventos

Interrupção de Hardware

- Sinal enviado à UCP por meio do barramento

Interrupção de Software

- Chamada de Sistema (System Call)
- Chamada ao Monitor (Monitor Call)

Tratamento de Interrupções

Ao receber uma interrupção a UCP

- Interrompe a tarefa atual
- Transfere execução para endereço fixo na memória
 - Este endereço contém o endereço inicial da rotina de atendimento da interrupção
- Retoma a tarefa interrompida

Quantidade de interrupções limitada e pré-definida

Vetor de Interrupções

Componentes do Sistema

Interação do usuário com o Sistema Operacional

- Em geral por meio de uma aplicação especial
- chamada shell

Núcleo (kernel)

- Software que contém componentes centrais do Sistema Operacional

Dentre os componentes mais comuns do sistema operacional encontra se

- Gerência de Processos
- Gerência da Memória
- Gerência de Arquivos
- Gerência Secundária
- Gerenciador de E/S
- Redes
- Sistema de Proteção
- Interpretador de Comandos
- Gerenciador de comunicação interprocessos (IPC)

Gerência de Processos

Um processo é um programa em execução

- Um processo precisa de certos recursos
 - tempo de CPU, memória, arquivos e dispositivos de E/S

O Gerenciamento de processos é responsável por

- (em conjunto com o SO)
- Criar e excluir processos
- Suspende e retomar processos
- Oferecer mecanismos para:
 - sincronismo de processos
 - comunicação entre processos

Programa é uma entidade passiva

Processo é uma entidade ativa

Processos - Mecanismos de Interrupção

Estabelece um sistema de sinalização de ocorrência de eventos

Proteção entre processos

- Impede situações de erros catastróficos

Processos - Definições

Processos são uma abstração para programa em execução

- Também chamado de tarefa, “task” ou “job”

Programa

- Entidade estática e permanente
- Invariável do ponto de vista do SO

Processo

- Entidade ativa e dinâmica
 - Possui estados que se alternam no tempo
 - Constituído de código, dados, contexto e recursos do sistema

Processos - Definições

Instâncias de programas

- Vários processos podem ser instâncias diferentes de um mesmo processo
- Mesmo código, porém dados e contexto de execução diferentes
- Processos são a forma com o SO vê programas

Processos executam

- Programas do usuário
- Programas de apoio ao SO

Estados de um Processo

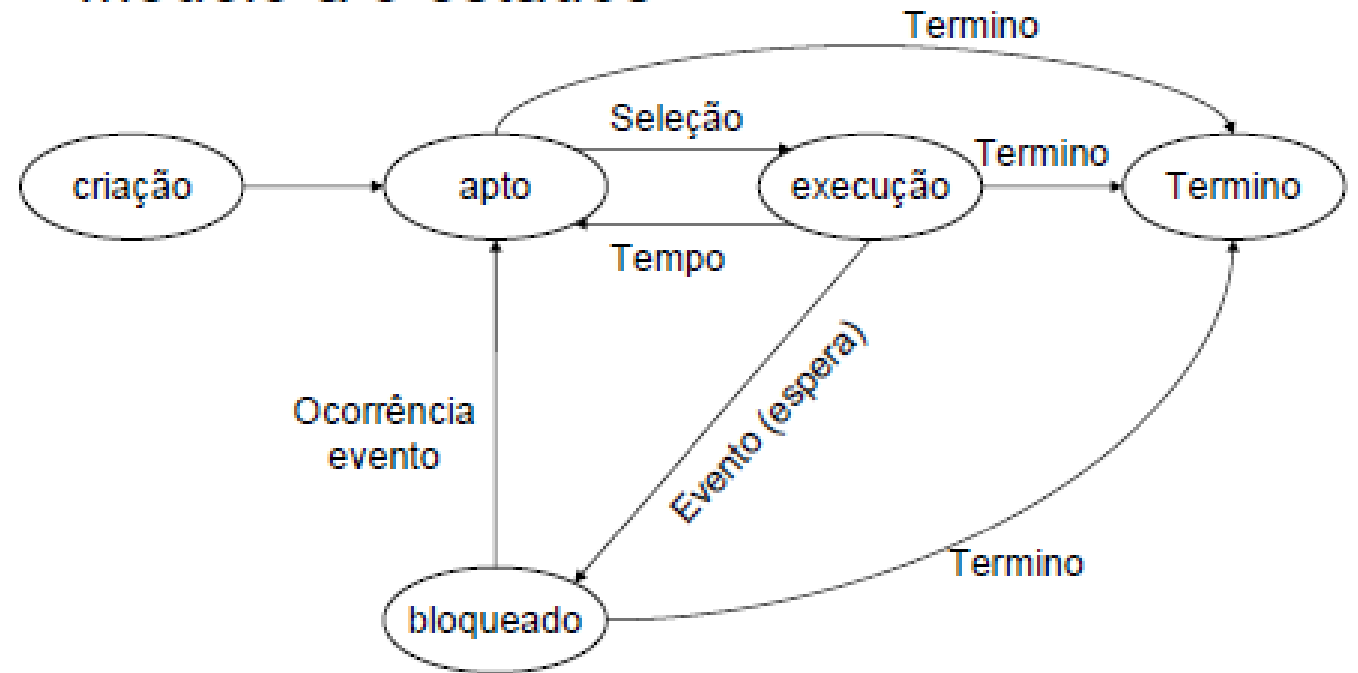
Só um processo pode ser executado a cada instante

- Manter uma fila de processos em espera de execução
- Selecionar periodicamente um novo processos da fila para execução

Prevenir que um processo monopolize o processador

Estados de um Processos

- Modelo a 5 estados

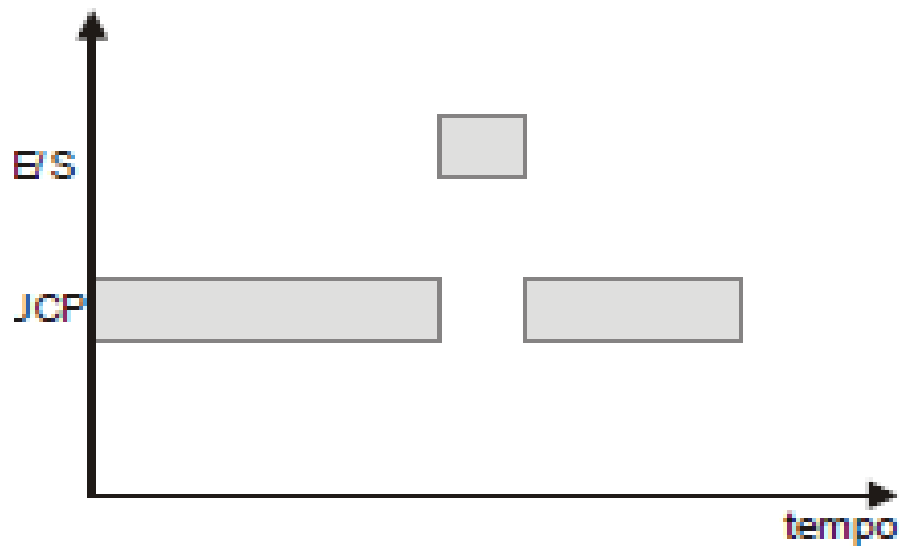


Transições de estado

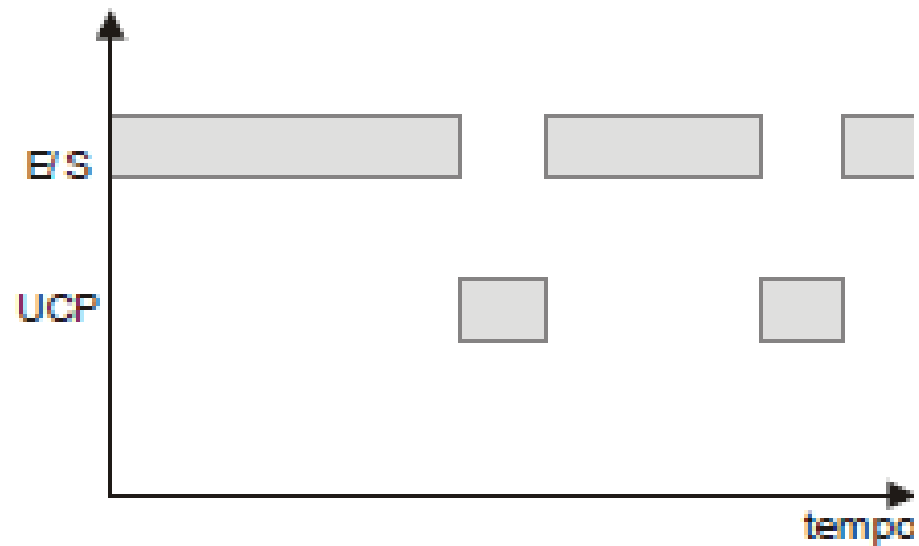
- Criação
 - Um novo processo é criado
 - Recursos são atribuído
 - Execução é necessariamente imediata
 - Apto → Execução
 - Processo é escolhido para execução
 - Depende da política de escalonamento
 - Execução → Término
 - Evento de termino

Processos em Execução

- CPU x E/S bound



(a) CPU-bound



(b) I/O-bound

Transições de estado

Execução → Apto

- Tempo máximo de execução sem interrupção foi atingida;
- Processo é “preemptado” pelo SO

Execução → Bloqueado

- Processo requisitou recurso/evento pelo qual deve esperar.

Bloqueado → Apto

- – Recurso disponível ou ocorrência do evento

Apto → Término e Bloqueado → Término

- Processo pai termina um processo filho
- Processo pai é terminado, e os processos filhos associados são também finalizados.

Suspensão de Processos

Processador é muito mais rápido que os dispositivos de E/S

- todos os processos em memória poderiam ficar em situação de espera.
- Ociosidade do processador
- Aumento de memória para acomodar mais processos

• SOLUÇÃO

- Liberar memória ocupado por processos bloqueados
- • Transferência dos processos para o disco (swapping)
- • Necessidade de novos estados
- Estado suspenso (processo no disco)

Gerência de Memória

Memória é um grande vetor de palavras ou bytes

- É um repositório de dados rapidamente acessíveis
- Cada elemento tem um com seu próprio endereço

Memória principal

- Dispositivo de armazenamento volátil
- Perde seu conteúdo em caso de falha de sistema

O Gerenciamento de memória é responsável por

- (em conjunto com o SO)
- Alocar e desalocar espaço de memória conforme a necessidade
- Controlar as partes da memória são usadas e por quem
- Decidir quais processos devem ser carregados
- Tratar do Swapping entre memória principal e memória secundária.
 - Transferência temporária de processos residentes na memória principal para memória secundária.

Memória



Memória principal:

também chamadas de memória real, são memórias que o processador pode endereçar diretamente, sem as quais o computador não pode funcionar.

Estas fornecem geralmente uma ponte para as secundárias, mas a sua função principal é a de conter a informação necessária para o processador num determinado momento;

Nesta categoria insere-se a memória RAM (volátil), memória ROM (não volátil), registradores e memórias cache.

Memória secundária

memórias que não podem ser endereçadas diretamente, a informação precisa ser carregada em memória principal antes de poder ser tratada pelo processador.

Não são estritamente necessárias para a operação do computador. São geralmente não-voláteis, permitindo guardar os dados permanentemente.

Incluem-se, nesta categoria, os discos rígidos, CDs, DVDs e disquetes.

memória terciária

A memória secundária não necessita de operações de montagem (inserção de uma mídia ou mídia em um dispositivo de leitura/gravação) para acessar os dados, como discos rígidos;

a memória terciária depende das operações de montagem, como discos ópticos e fitas magnéticas, entre outros

Técnicas de Alocação de Memória:

Alocação Contígua Simples;

Alocação Particionada Estática;

- Alocação Particionada Estática Absoluta;
- Alocação Particionada Estática Relocável;

Alocação Particionada Dinâmica.

Alocação Contígua Simples

Alocação implementada nos primeiros sistemas e ainda usada nos monoprogramáveis;

A Memória é dividida em duas áreas:

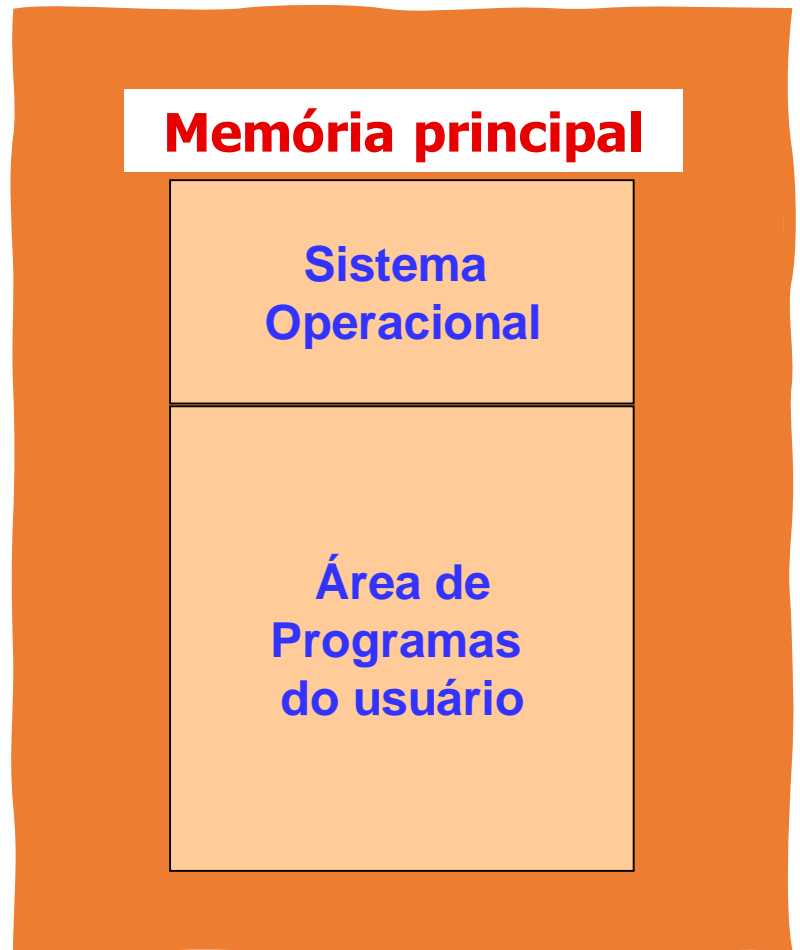
- Área do Sistema Operacional
- Área do Usuário

Um usuário não pode usar uma área maior do que a disponível;

Sem proteção:

- Um usuário pode acessar a área do Sistema Operacional.

Alocação Contígua Simples



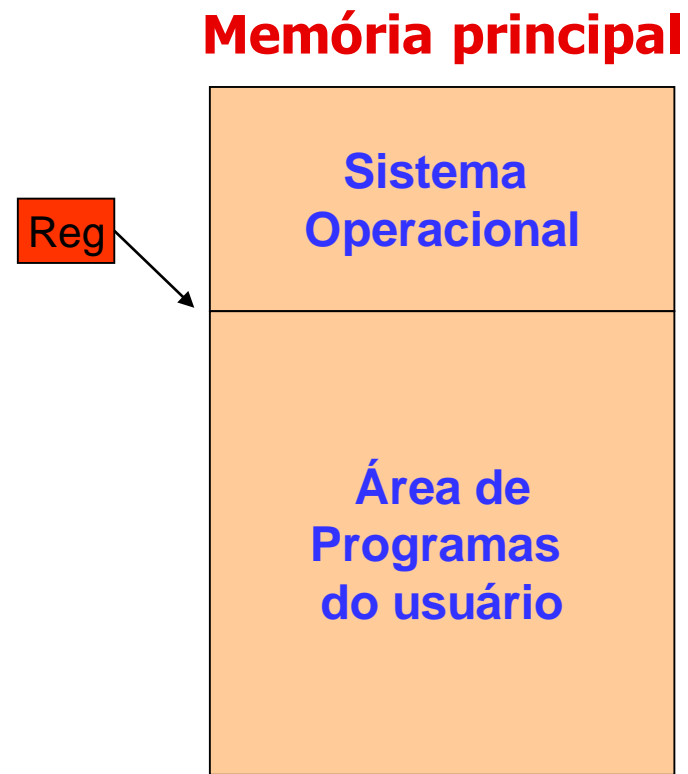
Alocação Contígua Simples

Registrador de proteção delimita as áreas do sistema operacional e do usuário;

Sistema verifica acessos à memória em relação ao endereço do registrador;

A forma de alocação era simples, mas não permitia utilização eficiente de processador e memória;

Alocação Contígua Simples



Alocação Contígua Simples

Programas de usuário limitados pelo tamanho da memória principal disponível.

Solução: Overlay

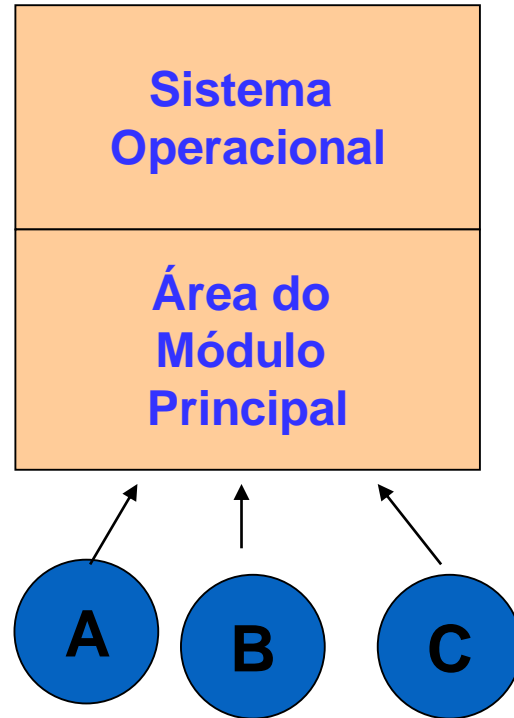
- Dividir o programa em módulos;
- Permitir execução independente de cada módulo, usando a mesma área de memória;

Área de Overlay

- Área de memória comum onde módulos compartilham mesmo espaço.

Alocação Contígua Simples

Memória principal



Alocação Particionada

Multiprogramação.

- Necessidade do uso da memória por vários usuários simultaneamente.

Ocupação mais eficiente do processador;

A memória foi dividida em pedaços de tamanho fixo chamados partições;

O tamanho de cada partição era estabelecido na inicialização do sistema;

Para alteração do particionamento, era necessário uma nova inicialização com uma nova configuração

Alocação Particionada Estática

Alocação Particionada Estática Absoluta:

- Compiladores gerando código absoluto;
 - Endereços relativos ao início da memória;
- Programas exclusivos para partições específicas.

Alocação Particionada Estática Relocável:

- Compiladores gerando código relocável;
 - Endereços relativos ao início da partição;
- Programas podem rodar em qualquer partição.

Algoritmos para localização em memória (Curiosidade)

O algoritmo de busca "best-first" usa a função heurística $F(n)=h(n)$ de procura ao nó de destino. Esta procura tenta expandir o nó que é mais próximo ao objetivo, acreditando numa condução rápida ao objetivo.

A heurística é aplicada globalmente, isto é, o caminho a ser seguido é selecionado entre todos os nodos abertos até o momento, O nodo aberto com a melhor nota é escolhido para a expansão

Algoritmos para localização em memória (Curiosidade)

- Worst-fit é um algoritmo utilizado na partição de disco rígido.
- Função
- O algoritmo worst-fit aloca o programa na pior partição, ou seja, aquela que deixa o maior espaço livre.
- Está técnica, apesar de aproveitar primeiro as partições maiores, acaba deixando espaços livres grandes o suficiente para que outros programas utilizem da memória, diminuindo ou retardando a fragmentação.

Algoritmos para localização em memória (Curiosidade)

- Next-fit é um algoritmo utilizado para partição de disco rígido.
- Processo
 - Algoritmo para partição dinâmica que inicia a busca a partir da posição da última alocação até encontrar o primeiro bloco, mais freqüentemente são alocados blocos de tamanho grande.
 - Grandes blocos são particionados em blocos menores e existe a necessidade de compactação quando não houver mais memória disponível.

Algoritmos para localização em memória (curiosidade)

First Fit é um algoritmo utilizado por sistemas operacionais para decidir em que local da memória um processo deve ser alocado.

Ele funciona escolhendo sempre o primeiro espaço vazio na memória que caiba o processo. Tem um tempo médio de espera entre processos relativamente baixo, principalmente se levado em conta sua facilidade de implementação.

O algoritmo Next Fit é um melhoramento do First Fit.

Existem os seguintes espaços livres: 11k, 3k, 19k, 18k, 7k, 8k, 13k, 15k. Se o First-Fit for utilizado, as solicitações 5k, 12k, 6k, ocupariam os espaços 11k, 19k, 18k respectivamente.