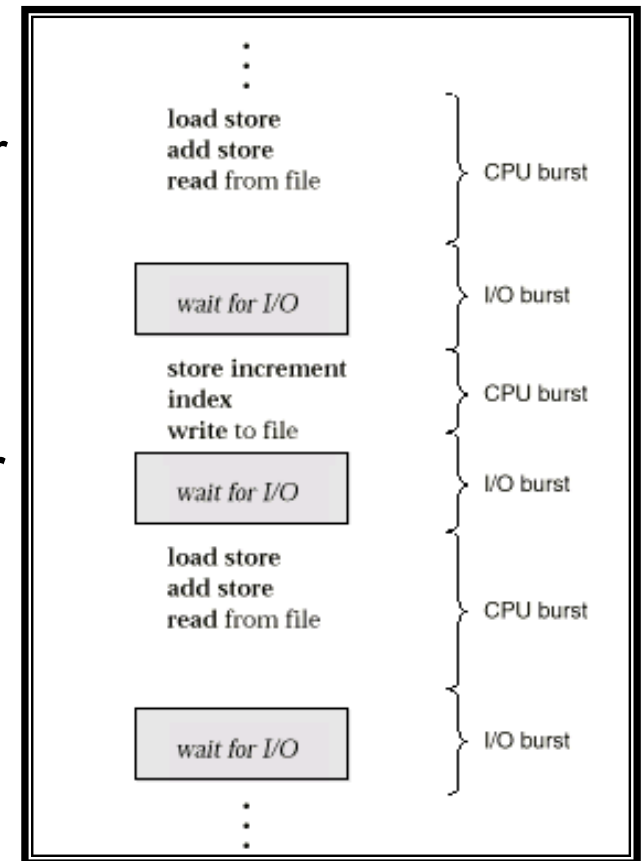


Escalonamento

- *Scheduling*
 - função fundamental do SO
 - quase todos os recursos do computador são submetidos ao scheduling antes do uso
- processos obedecem um ciclo que alterna processamento c/ espera por E/S
- distribuição de ciclos de CPU
- preocupação: troca de contexto



Situações típicas para execução do escalonador (1)

- Dependem se o escalonador é preemptivo ou não, se considera prioridades ou não ...
 - Sempre que a CPU estiver livre e houver processos aptos a executar
 - Criação e término de processos
 - Um processo de mais alta prioridade fica apto a executar
 - Interrupção de tempo
 - Ex.: processo executou toda a fatia

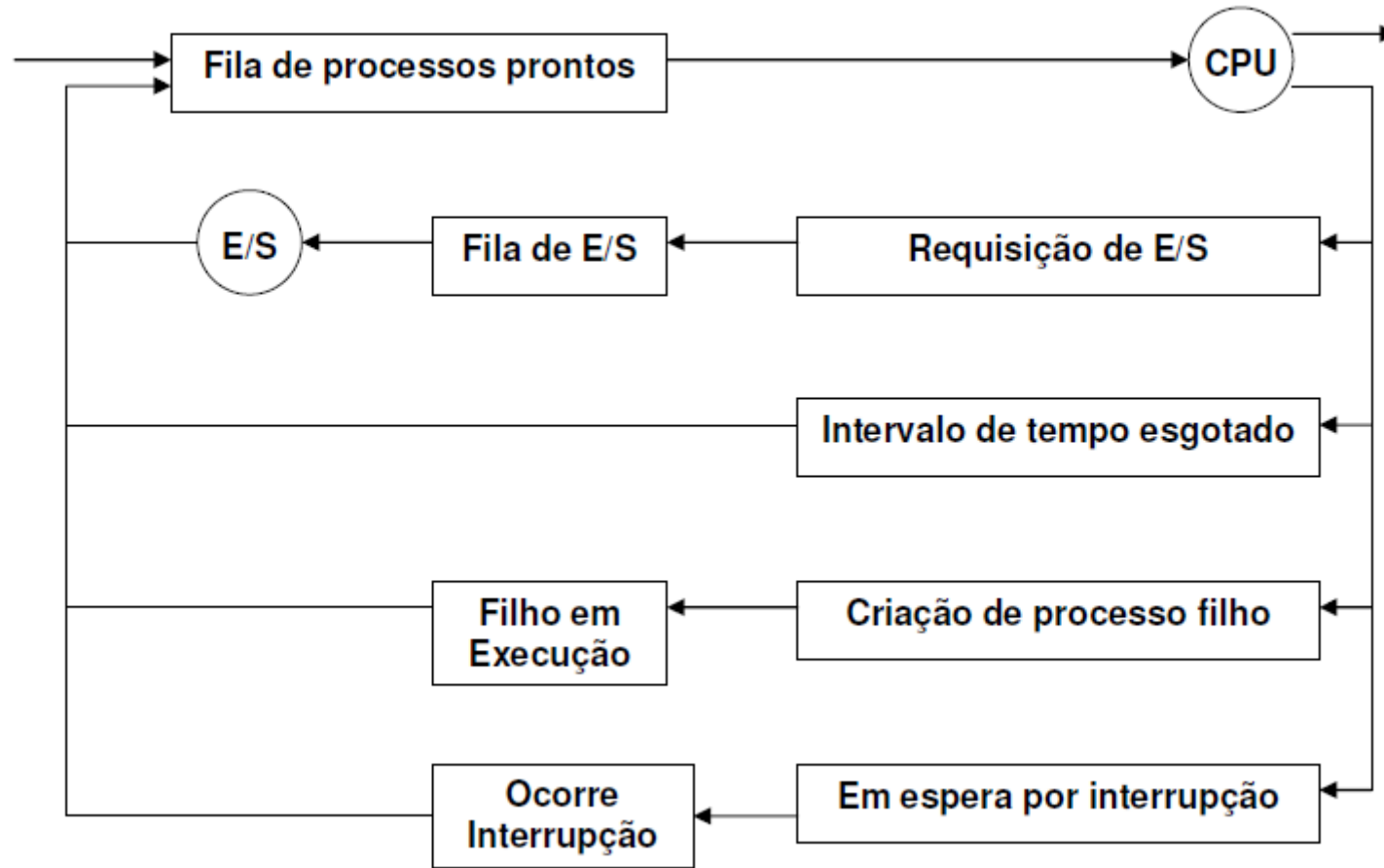
Situações típicas para execução do escalonador (2)

- Interrupção de dispositivos de E/S
- Interrupção por falta de página (ou segmento) em memória
 - Endereço acessado não está carregado na memória (está em disco)
- Interrupção por erros

Preempção

- Em SO, **preempção** é a capacidade de alterar a ordem de execução de (ou tirar de execução) um processo em detrimento de outro, de acordo com algum critério.
 - Preempção por tempo
 - término de fatia
 - tempo de ciclo de CPU
 - Preempção por prioridade

Diagrama de transição entre filas

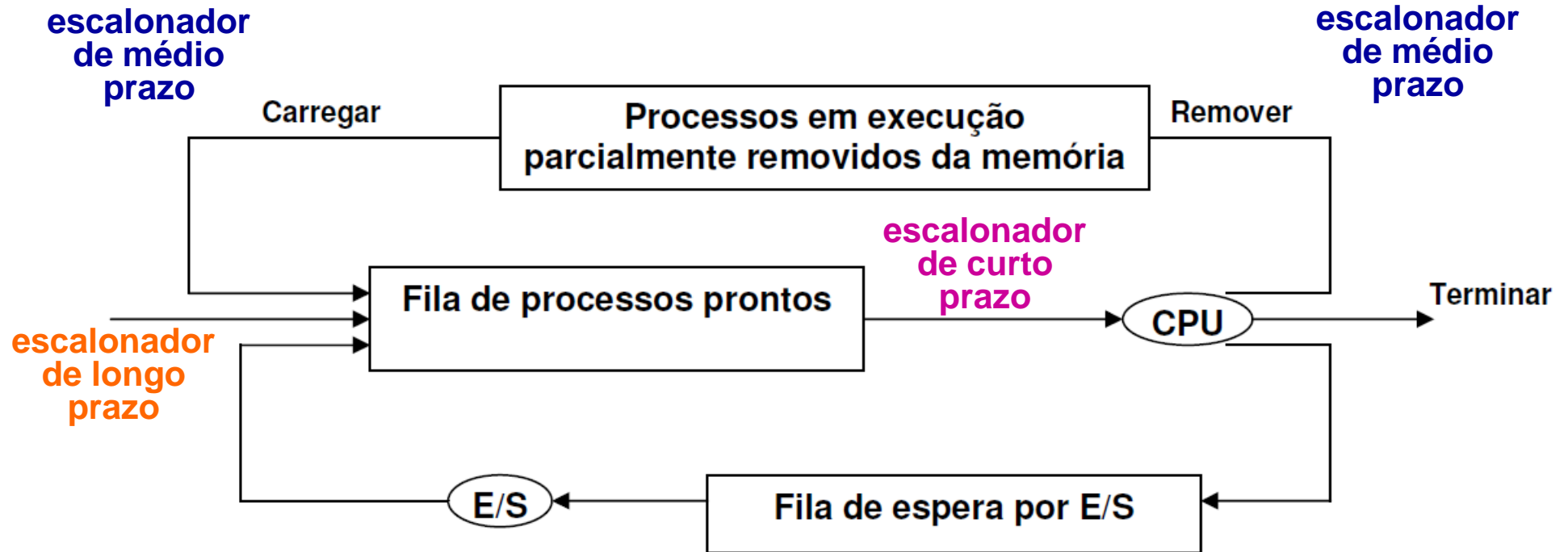


Escalonadores (1)

Curto prazo

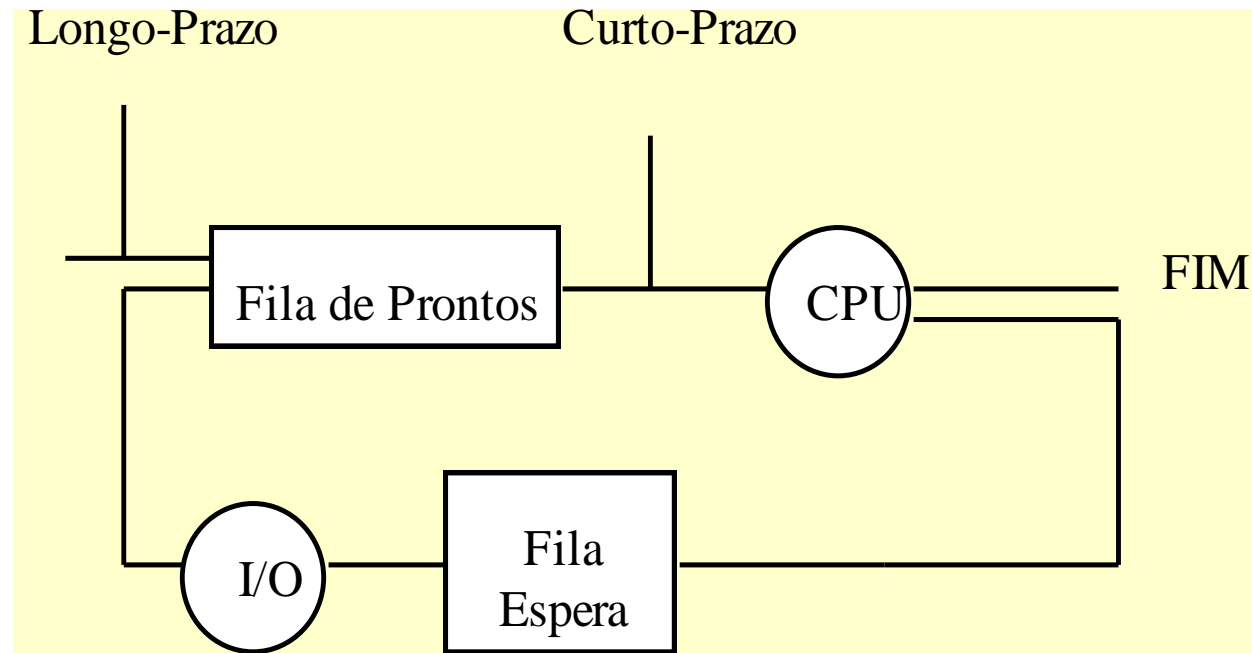
Longo prazo

Médio prazo



Escalonadores (2)

- curto prazo (**CPU scheduler/short-term scheduler**)
 - dispatcher
 - decide a alocação da CPU (seleciona o processo a executar e aloca CPU p/ ele)
 - MP → processador
 - invocado frequentemente (milisegundos)
 - deve ser rápido



Escalonadores (3)

- longo prazo (**job scheduler/long-term scheduler**)
 - gerencia a criação dos processos
 - seleciona processos p/ serem trazidos p/ fila de prontos (admissão ao sistema)
 - memória secundária → MP
 - pouco invocado (segundos, minutos)
 - pode ser lento
 - controla grau de multiprogramação (número de processos em memória)
 - em alguns sistemas, o escalonador de longo prazo pode não existir ou ser muito simples
 - Ex.: sistemas de tempo compartilhado frequentemente não tem um escalonador de longo prazo → colocam cada novo processo na memória, para o escalonador da CPU (curto prazo)
-

Escalonadores (4)

- médio prazo (**medium-term scheduler**)

- comum em sistemas com memória virtual

- gerencia a memória

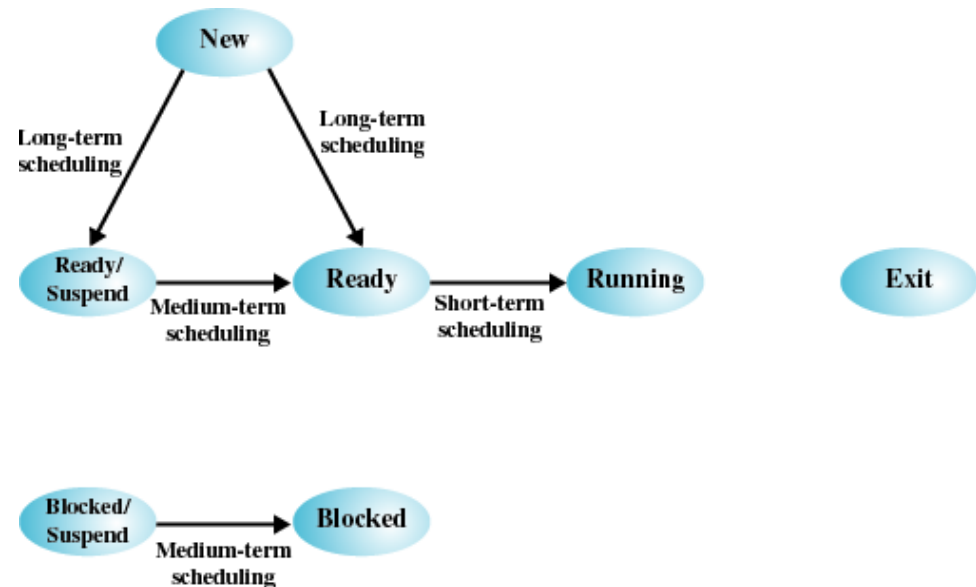
- escalonador de memória

- swapper

- associado ao swapping de processos

- objetivo: liberar espaço em memória

- remove processos que aguardam evento p/ área de swap
 - permite que novos processos sejam executados



Escalonador de longo prazo (*job scheduler*)

- Executa quando um processo é criado
- Determina quando um processo novo é admitido no sistema para processamento e disputa de recursos
- Controla o grau de multiprogramação do sistema
- Quanto maior o nro de processos admitidos, menor a porcentagem de tempo de CPU para cada um
- seleção criteriosa
 - *I/O-bound x CPU-bound*
 - diversos processos I/O-bound - fila ready quase sempre vazia
 - diversos processos CPU-bound - fila de dispositivos quase sempre vazia

Escalonador da CPU - *short-term scheduler*

- decisões de escalonamento ocorrem quando um processo troca do estado:
 1. “rodando” p/ “esperando”
 2. “rodando” p/ “pronto”
 3. “esperando” p/ “pronto”
 4. “rodando” p/ “terminated”
 - Escalonamento preemptivo
 - processo em execução pode ser suspenso p/ outro executar
 - ocorre nas 4 situações anteriores
 - Escalonamento não preemptivo
 - processo executa até o fim ou I/O
 - ocorre nas situações 1 e 4
-

CrITÉrios de escalonamento (1)

- Maximizar o uso da CPU
 - manter a CPU ocupada o máximo possível
- Maximizar o *throughput* (vazão do sistema)
 - *throughput* - n° de processos completados por unidade de tempo
- Minimizar o tempo de espera (*waiting time*)
 - soma dos tempos que um processo espera na fila de prontos

Critérios de escalonamento (2)

- Minimizar o tempo de *turnaround* (t de execução)
 - *turnaround* - tempo desde a submissão até o término da execução do processo
 - soma os tempos de espera p/ ocupar a memória, espera na fila de prontos, execução na CPU e realizando I/O
 - limitado pela velocidade do dispositivo de E/S
- Minimizar o tempo de resposta (*response time*)
 - tempo desde a submissão de um pedido até que a primeira resposta seja produzida

Dispatcher

- módulo que dá o controle da CPU ao processo selecionado pelo scheduler de curto prazo
- deve ser rápido
- envolve:
 - troca de contexto
 - troca p/ modo usuário
 - desvio p/ ponto exato no programa do usuário p/ reiniciá-lo
- latência de dispatcher
 - t entre parar um processo e iniciar outro