

**FACULDADE DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA  
LICENCIATURA EM ENGENHARIA INFORMÁTICA  
SISTEMAS OPERATIVOS E PROGRAMAÇÃO CONCORRENTE**

**TEMA: Despacho e escalonamento**

**Grupo Docente:**

- Eng<sup>o</sup>. Délcio Chadreca (Ms.C)
- Dr. Alfredo Covele (Ms.C)

# Tópicos da Aula

- ▶ Despacho
- ▶ Conceito de escalonamento
- ▶ Níveis de escalonamento
- ▶ Algoritmos de escalonamento
- ▶ Critérios de avaliação dos algoritmos de escalonamento

# Introdução

O escalonamento é muito fundamental para um computador porque ele faz a alocação e distribuição dos recursos do computador para os processos.

E a ordem de escalonamento é muito importante para a performance do computador. A escolha do algoritmo certo da ordem de escalonamento pode fazer toda diferença na *performance*. O Despacho também é muito importante porque ele dá controle do CPU ao processo que foi selecionado pelo escalonador.

# Despacho

Despacho é um mecanismo usado pelo sistema operativo no controle de mudança de contexto dos processos.

Além do escalonamento, o Despacho está também envolvido na função de agendamento do CPU, isto é, o Despacho da controlo do CPU ao processo que foi selecionado pelo escalonamento a curto prazo.

# Escalonamento

O escalonamento é um mecanismo usado pelos sistemas operacionais na gestão de concorrência entre os processos.

O escalonamento é muito fundamental para um computador, porque ele têm como função manter todos recursos computacionais ocupados e equilibrar o mesmo entre usuários e processos.

Ele faz a alocação e distribuição dos recursos para os processos.

# Níveis de Escalonamento

# Tipos de Escalonamento

O escalonamento é parte do Sistema Operativo, ele seleciona quais processos devem ser executados naquele momento e quais processos deverão ser executados a seguir. Ele têm a capacidade de pausar, mover e iniciar um processo.

Um Sistema Operativo pode ter até 3 tipos de escalonamento:

- ▶ Escalonamento a Longo prazo.
- ▶ Escalonamento a Medio prazo.
- ▶ Escalonamento a Curto prazo.

# Escalonamento a Longo prazo

O escalonamento a longo prazo, também chamado escalonamento de Admissão, decide que processos devem ser admitidos a lista de espera, se o processo for admitido ele é carregado para memória de execução.

As decisões que ele faz tem um efeito a longo termo no performance do computador. Os processos admitidos são retirados da memória secundária para a memória primária (de execução).

Este escalonamento têm como funções:

- ▶ Determinar se a execução de um processo é imediata;
- ▶ Decidir que processo é admitido a lista de espera;
- ▶ Garantir uma mistura de processos de E/S de dados e processos de CPU.



# Escalonamento a Medio prazo

O escalonamento a medio prazo tem a tarefa de remover temporariamente um processo da memoria principal para a memoria secundaria. Ele decide fazer essa troca com base nos seguintes critérios:

## **Actividade do processo**

- ▶ Se o processo não esta ativo há algum tempo ele troca de memória.

## **Prioridade do processo**

- ▶ Se um processo tem baixa prioridade ele troca de memória.

## **Memória ocupada pelo processo**

- ▶ Se um processo estiver a ocupar muita memoria primária, ele troca de memória.

# Escalonamento a Curto prazo

O escalonamento a Curto prazo, também chamado de escalonamento de CPU, tem função de decidir que processo está pronto para ser executado (ser alocado para CPU).

As decisões que ele fizer tem um efeito a curto prazo na performance do computador. Ele tem a capacidade de remover um processo em execução sem que o processo tenha chegado ao fim para alocar um outro processo. Este tipo de escalonamento têm como função:

- ▶ Aumentar o desempenho do sistema; e
- ▶ Decidir que processo executar a seguir.

# Filas de Agendamento

As filas de agendamento de processos ajudam a manter uma fila diferente para todos os estados dos processos. Todo o processo que contém o mesmo estado de execução é colocado na mesma fila e sempre que o estado do processo é modificado, seu PCB (Process Control Block) precisa ser modificado também.

Os três tipos de fila de agendamento são:

## **Filas de tarefas**

- ▶ Onde estão armazenados todos os processos do Sistema.

## **Fila Pronta**

- ▶ Onde estão os processos prontos para serem executados.

## **Fila de Dispositivos**

- ▶ Onde estão os processos bloqueados devido à ausência de dispositivo de E/S.

# Algoritmos de Escalonamento

Algoritmo de escalonamento são algoritmos usados para distribuição de recursos entre processos. Existindo vários algoritmos, e cada um com seu propósito, eles têm diferença significativas quando se trata de performance e utilização de espaço. Os algoritmos de escalonamento podem ser agrupados de acordo com a sua possibilidade de interromper um processo em execução:

- ▶ **Algoritmos preemptivos** – Não descartam a possibilidade de interromper um processo em execução.
- ▶ **Algoritmos não preemptivos** – Descartam a possibilidade de interromper um processo em execução.

# Tipos de Algoritmos de Escalonamento

Os algoritmos de escalonamento sejam eles preemptivos ou não preemptivos são destinados aos diferentes de sistemas nomeadamente:

- ▶ **Sistemas em lote ou *batch*** – Alta compatibilidade com os algoritmos de escalonamento que proporcionam longo intervalo de tempo aos processos e algoritmos de escalonamento não preemptivos pois não há um compromisso com o usuário no tempo de resposta rápida.

Exemplo: FIFO.

- ▶ **Sistemas interactivos** – Alta compatibilidade com algoritmos de escalonamentos preemptivos pois a um compromisso o usuário no tempo de resposta.

Exemplo: Round Robin.

- ▶ **Sistemas de tempo real** – Alta compatibilidade simultânea entre algoritmos preemptivos e não preemptivos pois geralmente são os processos processados e bloqueados rapidamente.

Exemplo: Escalonamento Garantido.

# Algoritmos de escalonamento

- ▶ FCFS (First-Come First-Served)
- ▶ SJF Não preemptivo (Shortest Job First)
- ▶ SJF Preemptivo (SRTF - Shortest Remaining Time First)
- ▶ Round-Robin
- ▶ Prioridade Fixa
- ▶ Escalonamento Garantido
- ▶ Filas múltiplas
- ▶ Lottery

# Critérios de Escalonamento

AT – Arrival Time

BT – Burst Time

CT – Completion Time

**TAT – Turnaround Time**

**WT – Waiting Time**

**RT – Response Time**

**TP – Throughput**

SL – Schedule Length

NP- Número de Processos

## Fórmulas

$$\mathbf{TAT} = \mathbf{CT} - \mathbf{AT}$$

$$\mathbf{WT} = \mathbf{TAT} - \mathbf{BT}$$

$$\mathbf{RT} = \mathbf{ReadyTime} - \mathbf{AT}$$

$$\mathbf{Throughput} = \mathbf{NP/SL}$$

# Critérios de Escalonamento

- ▶ **Utilização do Processador:** é desejável que o processador permaneça a maior parte do tempo ocupado.
- ▶ **Throughput:** representa o número de processos executados em um determinado instante de tempo.
- ▶ **Burst Time:** é o tempo que um processo leva no estado de execução durante o seu processamento.
- ▶ **Waiting Time:** é o tempo total que um processo permanece na fila de pronto durante seu processamento, aguardando para ser executado.



# Critérios de Escalonamento

- ▶ **Turnaround Time:** é o tempo que um processo leva desde a sua criação até seu término.
- ▶ **Response Time:** é o tempo decorrido entre uma requisição ao sistema ou à aplicação e o instante em que a resposta é exibida.
- ▶ **Schedule Length**
- ▶ **Completion Time**

# FCFS (First-Come First-Served)

O FCFS é o algoritmo mais simples ordenar os processos. Tendo uma lista de processos prontos para a execução, o primeiro que entrou na fila, será o primeiro a sair da mesma (a ser executado). Existindo um novo processo, o mesmo passa para o fim da fila.

Exemplo: Cinco processos, P1, P2, P3, P4 e P5, chegam num computador nos seguintes instantes 0, 5, 1, 6, 8 ms respectivamente. Eles têm seus tempos de processamento estimados em 8, 2, 7, 3 e 8 ms respectivamente.

- a) Faça o gráfico/diagrama Gantt.
- b) Qual é o tempo médio de espera?
- c) Qual é o Throughput?

	AT	BT	CT	TAT	WT	RT
P1	0	8	8	8	0	0
P2	5	2	17	12	10	10
P3	1	7	15	14	7	7
P4	6	3	20	14	11	11
P5	8	5	25	17	12	12

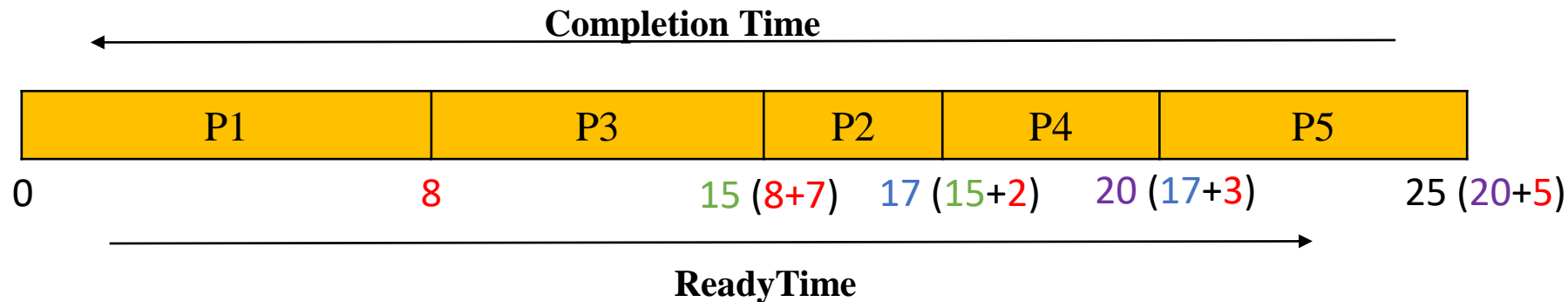
C) Qual é o Throughput?

$$TP = NP/SL$$

$$TP = 5/25 =$$

$$TP = 0.2 \text{ processos/ms.}$$

a) Faça o diagrama de Gantt.



b) Qual é o tempo médio de espera?

$$M_{WT} = (0+10+7+11+12)/5 = 8 \text{ ms/processo.}$$

Desafio: Como calcularia a produtividade e a utilização da CPU?

# SJF (Shortest Job First) – Não Preemptivo

Neste Algoritmo, os processos são executados por ordem crescente do tempo de execução, não se importando com o tempo de chegada de cada processo.

Exemplo: Cinco processos, P1, P2, P3, P4 e P5, chegam num computador nos seguintes instantes 0, 5, 1, 6, 8 ms respectivamente. Eles têm seus tempos de processamento estimados em 8, 2, 7, 3 e 8 ms respectivamente.

- a) Faça o gráfico Gantt.
- b) Qual é o tempo médio de espera?
- c) Qual é o Throughput?

	AT	BT	CT	TAT	WT	RT
P1	0	8	8	8	0	0
P2	5	2	10	5	3	3
P3	1	7	25	24	17	17
P4	6	3	13	7	4	4
P5	8	5	18	10	5	5

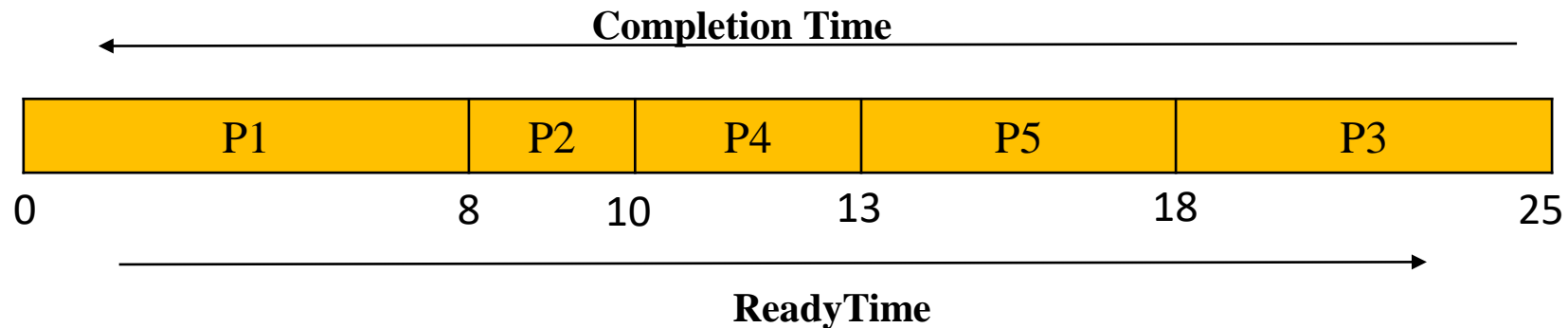
C) Qual é o Throughput?

$$TP = NP/SL$$

$$TP = 5/25 =$$

$$TP = 0.2 \text{ processos/ms.}$$

a) Faça o diagrama de Gantt.



b) Qual é o tempo médio de espera?

$$M_{WT} = (0+3+17+4+5)/5 = 5.8 \text{ ms/processo.}$$

b) Qual é o tempo médio de resposta?

$$M_{RT} = (0+3+17+4+5)/5 = 5.8 \text{ ms/processo.}$$

# SJF Preemptivo (SRTF - Shortest Remaining Time First)

Uma comparação é feita entre o Burst time de um novo processo novo existente com o tempo restante do tempo do processo em execução, caso o Burst Time do novo seja menor que o tempo restante do processo em execução, o novo processo toma a posse do processador, caso não, o processo em execução continua.

Um Sistema Operativo usa o algoritmo SJF para o escalonamento preemptivo de processos. Supondo que processos cheguem na ordem: P1 , P2 , P3, P4 nos instantes 0, 2, 4, 5 ms respectivamente. E com tempos de processamento estimados em 7, 4, 1 e 4 ms respectivamente.

- a) Faça o gráfico Gantt.
- b) Qual é o tempo médio de espera?
- c) Qual é o Throughput?

	AT	BT	CT	TAT	WT	RT
P1	0	7	16	16	9	0
P2	2	4	7	5	1	0
P3	4	1	5	1	0	0
P4	5	4	11	6	2	2

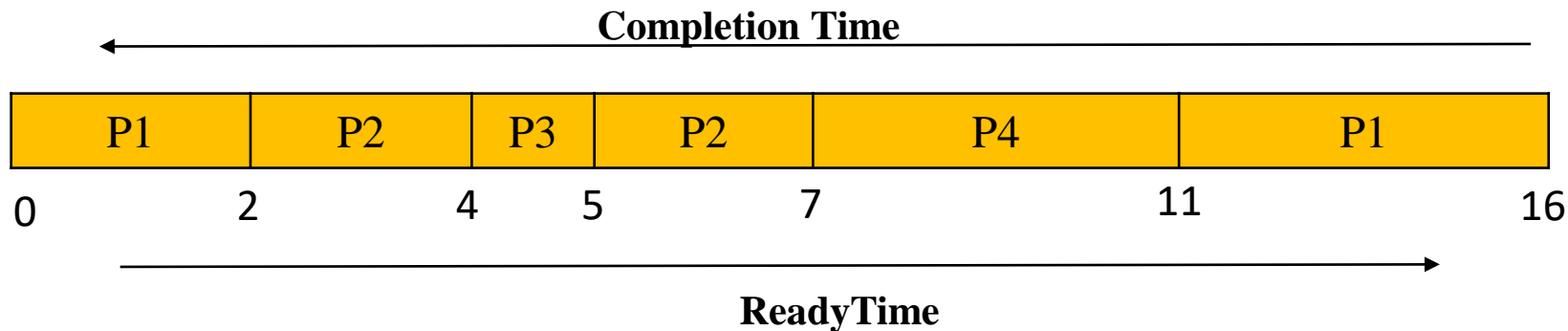
C) Qual é o Throughput?

$$\mathbf{TP} = \mathbf{NP/SL}$$

$$\mathbf{TP} = 5/16 =$$

$$\mathbf{TP} = 0.3125 \text{ processos/ms.}$$

a) Faça o diagrama de Gantt.



b) Qual é o tempo médio de espera?

$$\mathbf{M_{WT}} = (9+1+0+2)/4 = 3 \text{ ms/processo.}$$

b) Qual é o tempo médio de resposta?

$$\mathbf{M_{RT}} = (0+0+0+2)/5 = 0.4 \text{ ms/processo.}$$

# Round-Robin

O conceito quantum é um factor essencial, pois atribui-se o *quantum* (tempo limite de execução) a todos os processos da fila, sendo que, dependendo do factor de prioridade, espera-se que o primeiro (bem como todos os outros) processo que entra em execução, termine dentro do *quantum*, caso isso não seja possível, é interrompido esperando a próxima oportunidade de execução, e dando oportunidade ao outro processo.



	AT	BT	CT	TAT	WT	RT
P1	0	8				
P2	5	2				
P3	1	7				
P4	6	3				
P5	8	5				

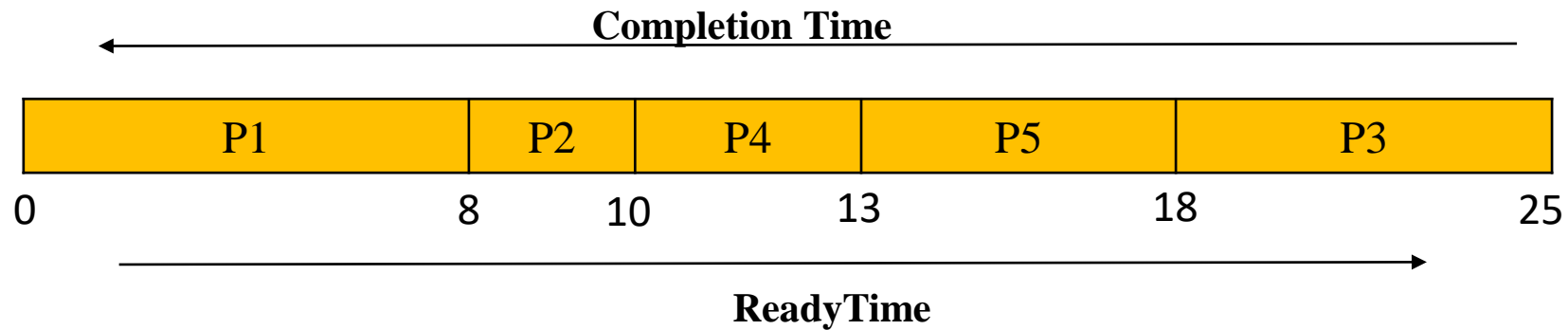
C) Qual é o Throughput?

$$TP = NP/SL$$

$$TP = =$$

$$TP = \text{processos/ms.}$$

a) Faça o diagrama de Gantt.



b) Qual é o tempo médio de espera?

$$M_{WT} = ()/ = \text{ms/processo.}$$

b) Qual é o tempo médio de resposta?

$$M_{RT} = ()/ = \text{ms/processo.}$$

# Questões de reflexão

1. Qual é a relação entre um programa e um processo?
2. O que significa um processo sofrer preempção?
3. Defina *turnaround time*, *waiting time* e *throughput* em suas palavras?
4. Em poucas palavras explique o que acontece no Algoritmo de Escalonamento Shortest Job First (SJF), Shortest Remaining Time First (SRTF), por Prioridade, Round-Robin.
5. Por que tanto o escalonamento FCFS quanto o SJF não são utilizados em sistemas de *time-sharing*?

**OBRIGADO !!!**