

# Sistemas en Tiempo Real

---

## Asignación y planificación de tareas

Francisco Andrés Candelas Herías

fcandela@dfists.ua.es



Universitat d'Alacant  
Universidad de Alicante

## Objetivos de la clase

---

2

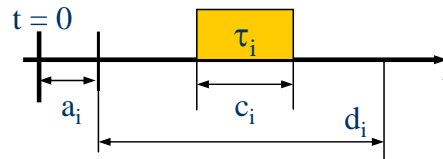
- Diferenciar entre planificación estática y dinámica.
- Estudiar como se abordan los problemas de asignación y planificación.
- Conocer los métodos básicos de asignación espacial.
- Conocer algunos algoritmos clásicos de planificación.

- Definiciones
- Asignación espacial
- Planificación temporal
- Algoritmos de planificación temporal
- Resumen

- Tarea: unidad de ejecución básica para la asignación y planificación, obtenida de la etapa de particionado software.
- Las tareas pueden ser...
  - Esporádicas o periódicas.
  - Interrumpibles o no interrumpibles.
  - Críticas o no críticas.
- Existen relaciones entre las tareas:
  - Relaciones de precedencia.
  - Relaciones de prioridad (orden de interrupción).
  - Relaciones de exclusión en el acceso a recursos

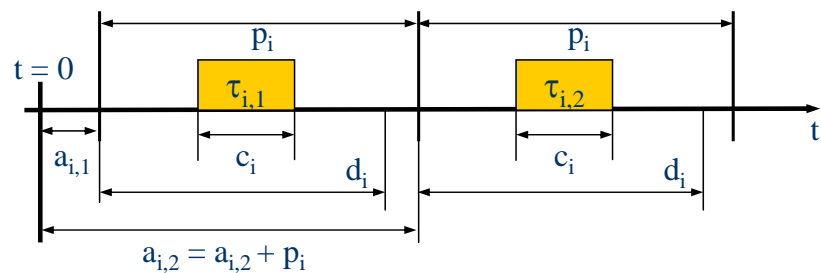
- Parámetros de las tareas:
  - Tareas esporádicas.

$$\tau_i = (a_i, c_i, d_i)$$



- Tareas periódicas.

$$\tau_i = (a_i, c_i, d_i, p_i)$$

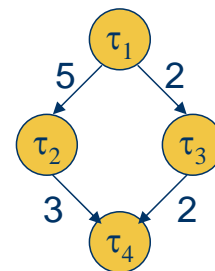


- Las relaciones de precedencia se pueden expresar con un grafo dirigido sin ciclos (DAG).

$$G = (T, E)$$

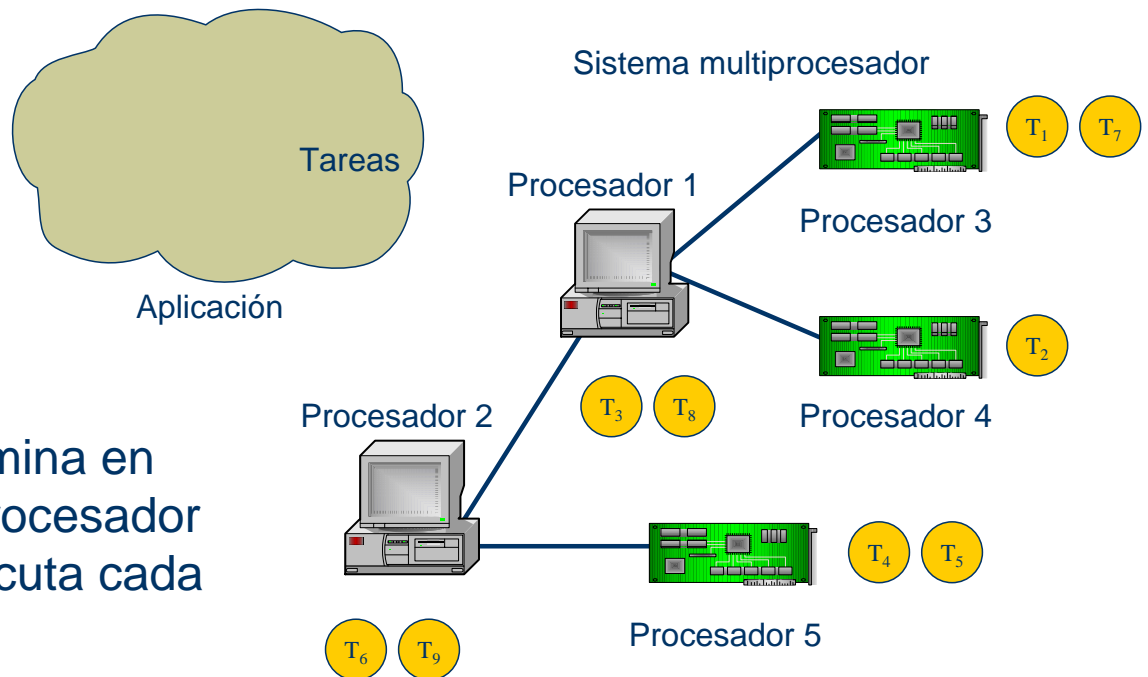
$$T = \{\tau_1, \tau_2, \dots, \tau_N\}, \quad \tau_i = (a_i, c_i, d_i)$$

$$E = \{(\tau_i, \tau_j, c_{i,j}) / \tau_i R_{\text{prec}} \tau_j \text{ con coste } c_{i,j}\}$$

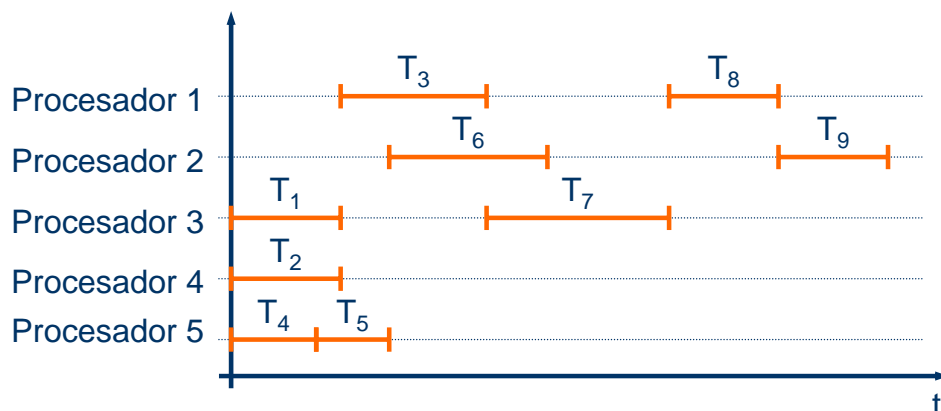


## Asignación espacial de las tareas (allocation):

- Determina en que procesador se ejecuta cada tarea.

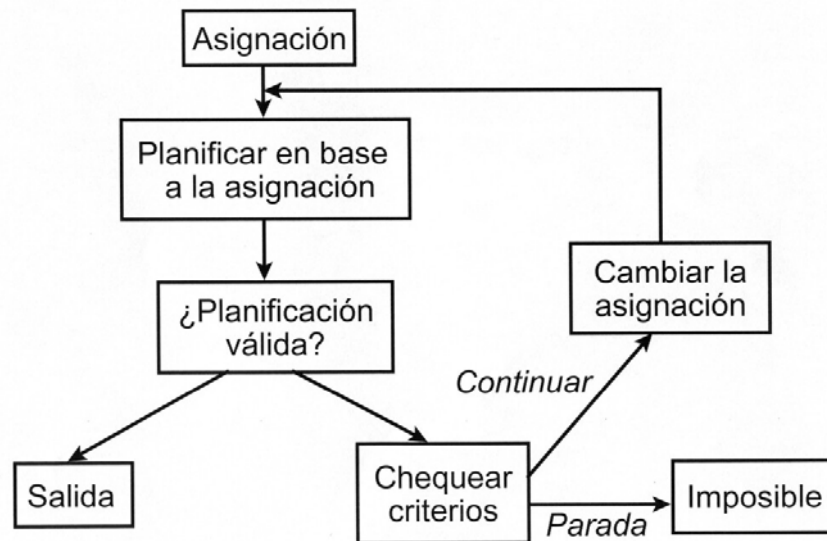


- ## Planificación temporal de las tareas (scheduling):
- Ordenar en el tiempo la ejecución de las tareas asignadas a cada procesador.



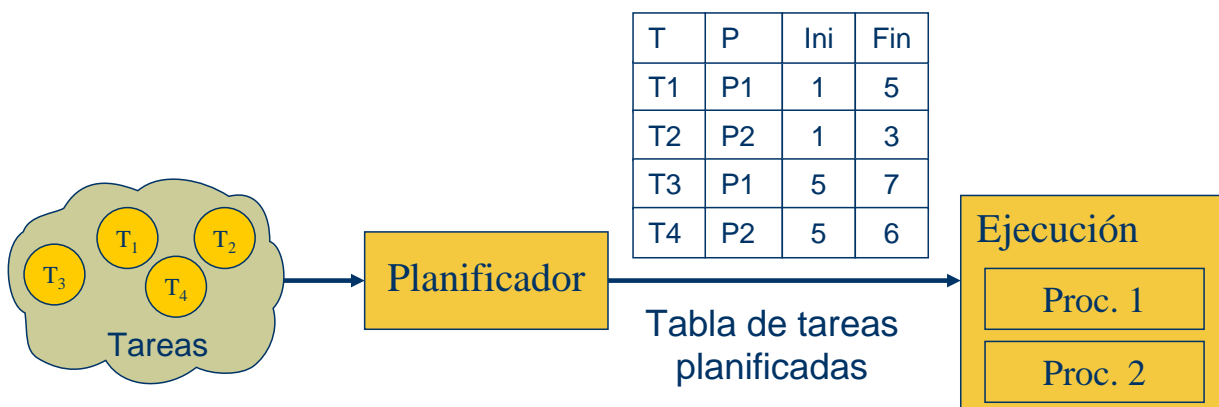
- Se suele hablar de planificación para referirse a la asignación y a la planificación temporal en conjunto.

## Desarrollo de la planificación:



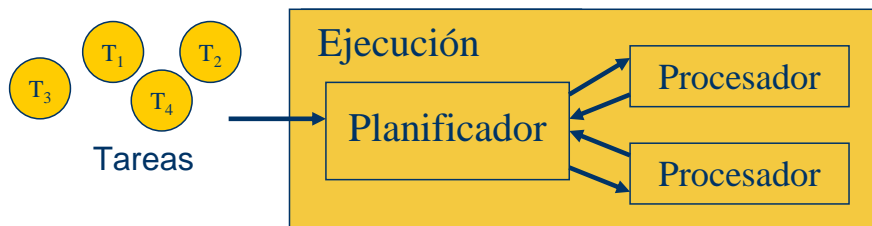
- También hay algoritmos que realizan la asignación y la planificación de forma simultánea.

## Planificación estática (off-line)



- Las características de las tareas se conocen a priori.
- Para sistemas bien conocidos.
- Ventajas: bajo coste, muy predecible, rápida implementación.
- Inconvenientes: no considera situaciones imprevistas ni posibles cambios en el sistema

## Planificación dinámica (on-line o run-time)



- No se requiere conocer bien las características de las tareas.
- Para sistemas con llegadas impredecibles o con incertidumbre en sus parámetros.
- Ventajas: Adecuada para sistemas que cambian, flexible.
- Inconvenientes: coste computacional en tiempo de ejecución.

- Definiciones
- Asignación espacial
- Planificación temporal
- Algoritmos de planificación temporal
- Resumen

- Objetivo: obtener la distribución *óptima* de las tareas en los procesadores disponibles.
- Tienen sentido en un sistema multiprocesador o multicomputador.
- Se pueden considerar...
  - Que todos los procesadores son del mismo tipo, y todas las tareas también.
  - Que hay procesadores de diferentes tipos, y cada uno ejecuta un tipo de tareas determinado.

- Criterios de optimización:
  - Minimizar el tiempo de respuesta total del sistema (minimizar coste de ejecución).
  - Minimizar la comunicación entre procesadores:
    - Menos datos intercambiados.
    - Menos retrasos en la propagación de datos.
  - Obtener una distribución de carga balanceada entre todos los procesadores.
  - Minimizar el número de recursos compartidos.
- Pero un algoritmo genérico de asignación es un problema NP-completo.

- Soluciones prácticas:

- Asignación óptima solo en casos especiales. Se requiere imponer restricciones como...

- Un número limitado de tareas.
- Se dispone de todos los procesadores necesarios.
- Las tareas no tienen relaciones de precedencia.
- No hay costes de comunicación entre tareas.
- Las tareas son de un tipo determinado...

- Soluciones óptimas locales.

- Soluciones subóptimas.

} No son óptimas, pero son aceptables para aplicaciones concretas.

- Hay muchos algoritmos diseñados a medida para aplicaciones determinadas.

- Soluciones óptimas locales:

- Métodos interactivos.

- Se hacen pequeños cambios en la asignación inicial o actual, se evalúa la asignación, y si no es válida se repite el proceso.
- Se puede usar algoritmos genéticos.

- Programación matemática.

- Se define una función objetivo a minimizar, un conjunto de restricciones (p.e. relaciones entre tareas) y se aplica un método de optimización (técnicas heurísticas).

- Búsqueda del espacio de estados.

- Se considera un árbol de búsqueda con las posibles asignaciones, se define una función de coste, y se busca la mejor opción. Se puede “podar” soluciones no aceptables.

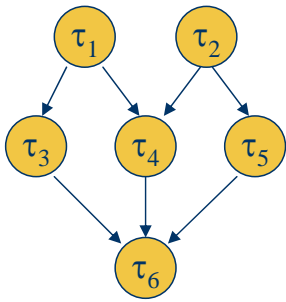


- Soluciones subóptimas:
  - Pasos:
    - Se considera una lista de tareas con prioridades para ser asignadas a los procesadores.
    - Se asigna las tareas más prioritarias en los procesadores más adecuados.
  - Para asignar las prioridades se puede considerar....
    - Los tiempos de ejecución de las tareas, las relaciones de precedencia, el número de sucesores de una tarea, el camino críticos de una tarea en el DAG, o varios factores juntos.
  - Para seleccionar un procesador se puede considerar...
    - Costes de comunicación, el procesador con menos carga, el procesador que acabará antes la tarea...

- Definiciones
- Asignación espacial
- Planificación temporal
- Algoritmos de planificación temporal
- Resumen



## ▪ Ejemplo sencillo:



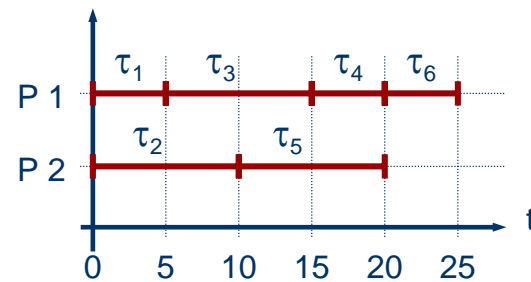
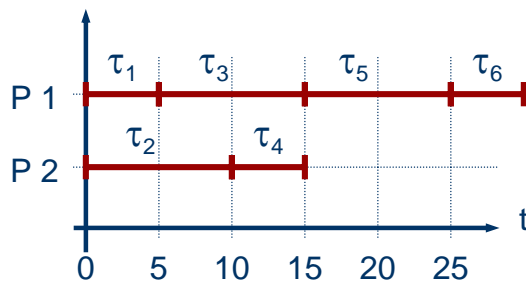
$$\tau_i = (a_i, c_j, d_i)$$

$$\tau_1 = (0, 5, 10) \quad \tau_4 = (5, 5, 15)$$

$$\tau_2 = (0, 10, 15) \quad \tau_5 = (5, 10, 20)$$

$$\tau_3 = (0, 10, 15) \quad \tau_6 = (10, 5, 15)$$

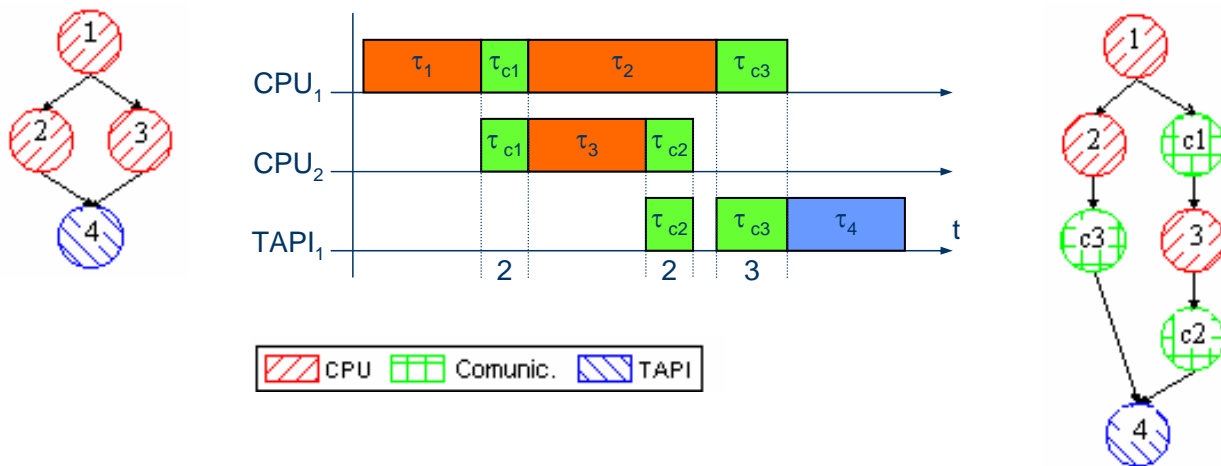
- Dos procesadores.
- No pre-emptive.
- Costes de comunicación nulos.



## ▪ Costes de comunicación:

- Cuentan cuando dos tareas relacionadas que necesitan intercambiar datos se ejecutan en diferentes procesadores.
- Muchos algoritmos los consideran nulos.
- Las comunicaciones se pueden considerar como nuevas tareas no interrumpibles a la hora de planificar.

- Costes de comunicación:
  - Ejemplo:



## Contenidos

- Definiciones
- Asignación espacial
- Planificación temporal
- Algoritmos de planificación temporal
- Resumen

- Un algoritmo de planificación temporal genérico también es un problema NP-completo.
- En la práctica el problema se aborda para aplicaciones concretas, considerando heurísticas y restricciones.
- Algoritmos monoprocesador básicos:
  - Sin relaciones de precedencia ni exclusión:
    - RM, DS, EDF.
  - Hay versiones con relaciones de precedencia y exclusión.

- Ejemplos de algoritmos multiprocesador:
  - Para determinadas aplicaciones, con restricciones:
    - Balanceado de carga.
    - Next-Fit, basado en RM.
    - Bink-Packet, basado en EDF.
    - MOS: búsqueda en un árbol con poda.

- RM (Rate Monotonic):
  - Tareas periódicas con periodo igual al deadline ( $d_i=p_i$ ), sin relaciones de precedencia, y no interrumpibles.
  - Cada tarea tiene una prioridad fija, inversamente proporcional a su periodo.
  - Con las condiciones anteriores es óptimo.
  - Hay muchas versiones de este algoritmo para otras restricciones.
- DS (Deferred Server):
  - Basado en RM.
  - Permite tareas periódicas y no periódicas simultáneamente.

- EDF (Earliest Deadline First):
  - Tareas interrumpibles, pero sin independientes.
  - La prioridad de una tarea depende del plazo de respuesta de forma dinámica: la tarea cuyo deadline llega antes es la más prioritaria.
  - Es óptimo con las restricciones anteriores.

- **Balanceado de carga:**
  - Reparte las tareas entre los procesadores de forma equitativa. Las tareas se asignan al procesador menos ocupado.
  - Puede considerar tareas interrumpibles.
  - Para seguridad ante fallos, las tareas se pueden duplicar, y tener copias en varios procesadores.
- **Next Fit (basado en RM):**
  - Tareas interrumpibles y periódicas, cuyo único recurso requerido es el procesador. Procesadores idénticos.
  - Agrupa las tareas en clases, y a asigna a cada una clase de tareas a ejecutar.

- **Bin-Packing (basado en EDF):**
  - Tareas periódicas con periodo igual al deadline ( $d_i=p_i$ ), sin relaciones de precedencia, aunque interrumpibles.
  - El único recurso requerido es el procesador. Los procesadores son idénticos.
  - La planificación trata de evitar que la utilización ( $U$ ) de un procesador 1 o menor.
- **MOS (Miopic Offline Scheduling):**
  - Tareas no interrumpibles que usan varios tipos de recursos.
  - Emplea una heurística para buscar una buena solución en un árbol con las posibles soluciones.

$$U = \sum_{i=1}^N \frac{c_i}{p_i}$$

- Definiciones
- Asignación espacial
- Planificación temporal
- Algoritmos de planificación temporal
- Resumen

- La planificación puede ser estática o dinámica.
- La planificación consta de una asignación espacial y una planificación temporal.
- Los problemas de asignación y planificación son NP-completos en general.
- Existen diferentes algoritmos para casos concretos, cada uno con distintas restricciones.
- Hay soluciones óptimas para casos muy concretos.