

ISE-T1-Apuntes.pdf



marinamuca01



Ingeniería de Servidores



3º Grado en Ingeniería Informática



**Escuela Técnica Superior de Ingenierías Informática y de
Telecomunicación
Universidad de Granada**

TEMA 1

Introducción a la ingeniería de servidores

Concepto de servidor

Sistema informático (S.I.)

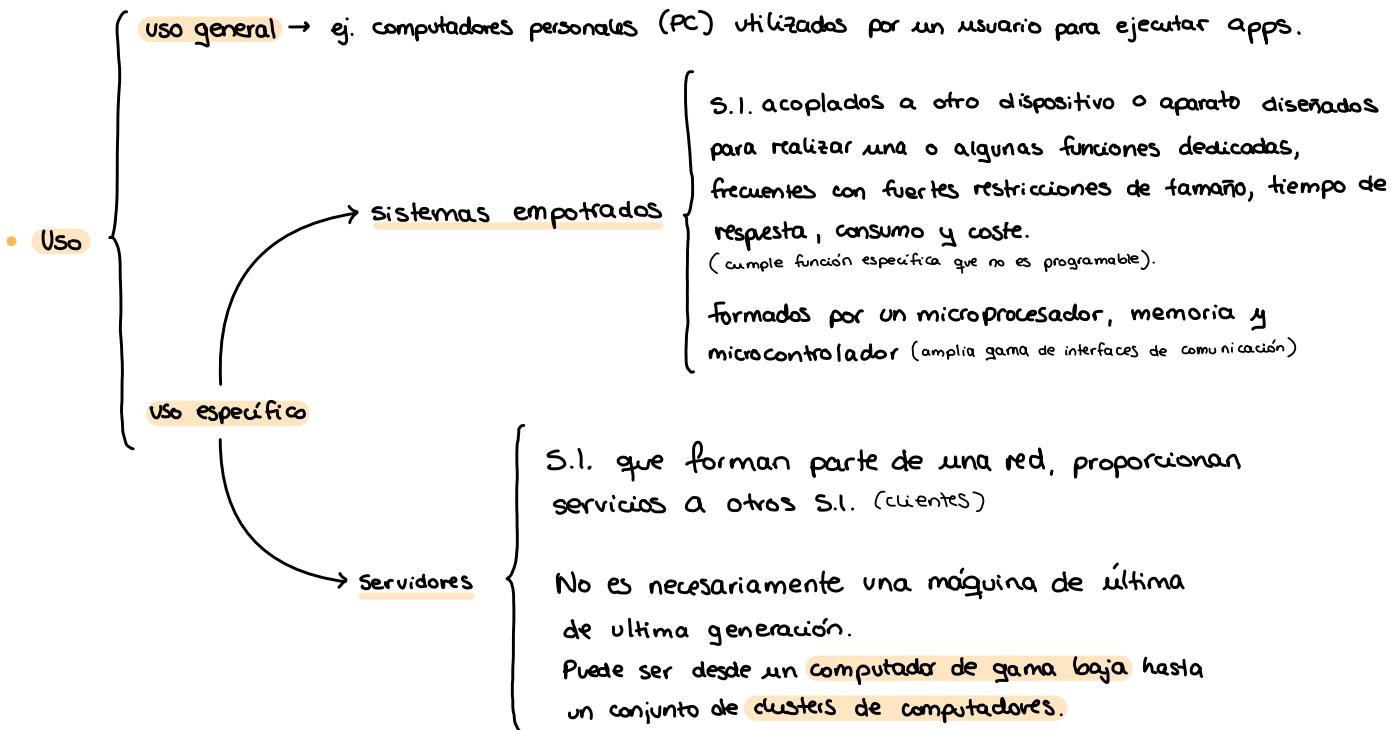
Conjunto de elementos **Hardware**, **Software** y **Peopleware** interrelacionados entre sí que permiten **obtener**, **almacenar** y **procesar información**.

- **Hardware** → componentes físicos del S.I. (CPU, memoria, cables, almacenamiento).
- **Software** → componentes lógicos del S.I. (S.O. y apps)
- **Peopleware** → conjunto de recursos humanos. (Personal técnico que configura y mantiene el S.I. usuarios que lo utilizan).

Clasificación de Sistemas Informáticos

Pueden clasificarse según numerosos criterios:

- **Paralelismo** de la arquitectura de procesamiento



TIPOS:

Web → almacena documentos HTML, imágenes, archivos de texto, ... y distribuye este contenido a clientes que lo soliciten en la red.

de archivos → permite el acceso remoto a archivos almacenados en él o directamente accesibles por este.

de base de datos → provee servicios de base de datos a otros programas u otras computadoras.

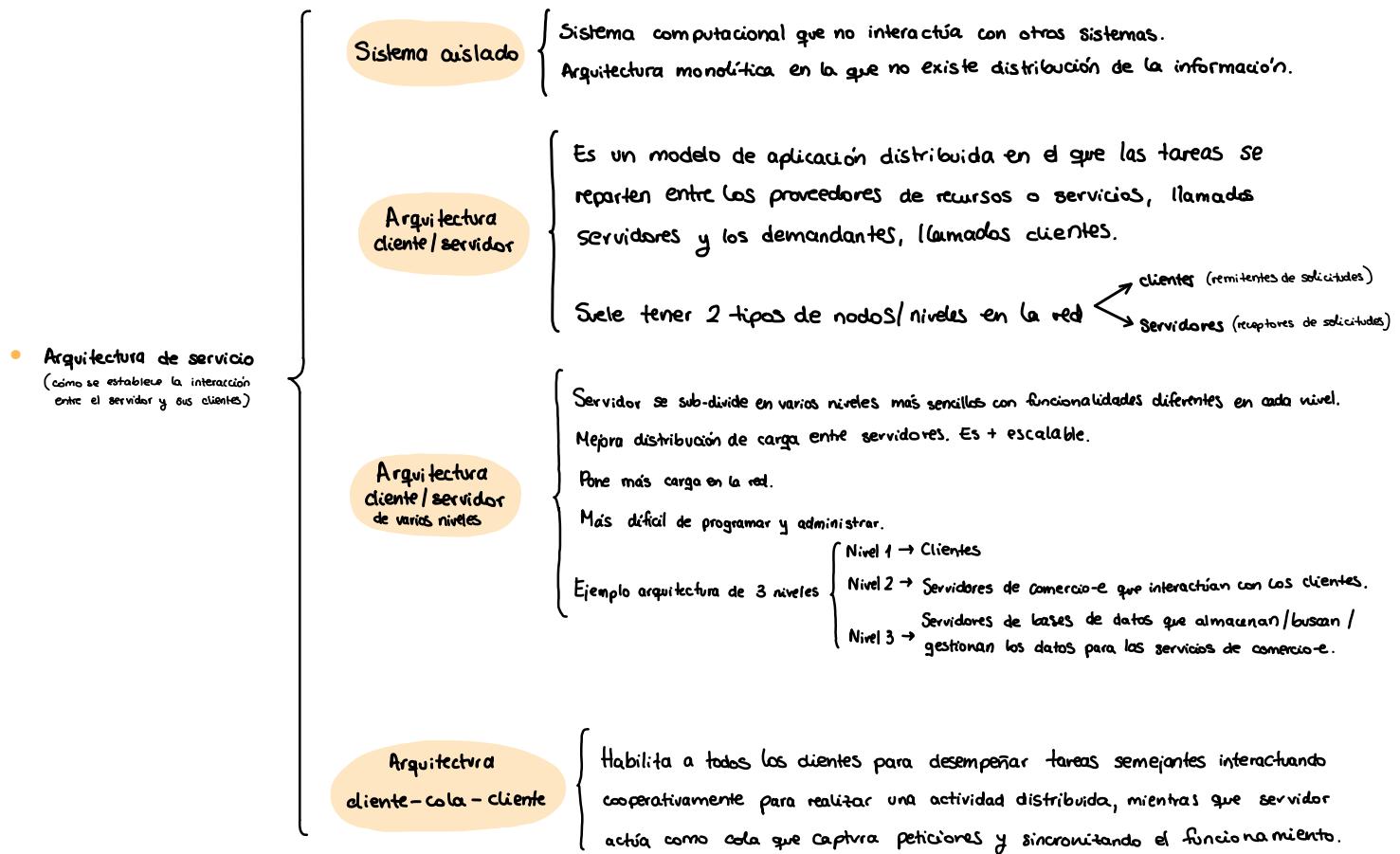
de comercio-e → cumple o procesa transacciones comerciales. Valida el cliente y genera un pedido al servidor de bases de datos.

de correo-e → almacena, envía, recibe, re-enruta y realiza otras operaciones relacionadas con e-mail para los clientes de la red.

DHCP → asigna dinámicamente una dirección IP y otros parámetros de configuración de red a cada dispositivo

DNS → devuelve la IP asociada a un nombre de dominio.

de impresión → controla una o más impresoras y acepta trabajos de impresión de otros clientes de la red.



Fundamentos de ingeniería de Servidores

Diseño, configuración y evaluación de un servidor

Recursos físicos, lógicos y humanos:

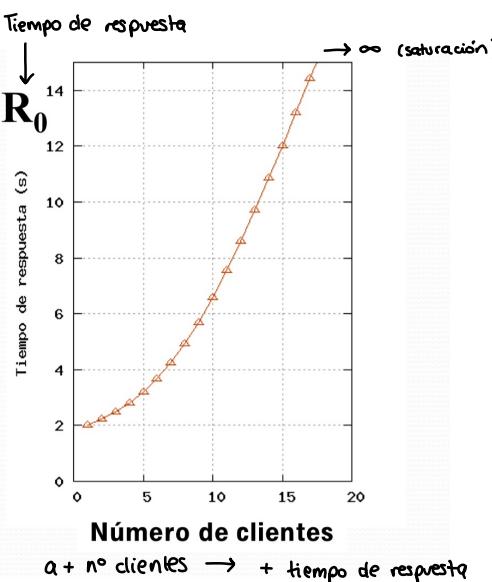
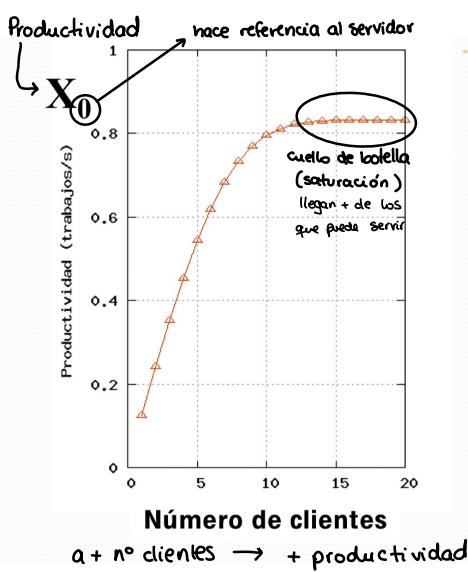
- Placa Base
- SO
- Memoria
- Apps
- Microprocesador
- Conexiones de red
- Fuentes de alimentación
- Cableado
- Periféricos
- Refrigeración ...

Requisitos funcionales (no funcionales de FIS)

- | | | |
|-----------------|------------------|-----------------|
| • Prestaciones | • Disponibilidad | • Fiabilidad |
| • Seguridad | • Extensibilidad | • Escalabilidad |
| • Mantenimiento | • Coste | |

Prestaciones (Performance)

- Medida / cuantificación de la velocidad con que se realiza una determinada carga o cantidad de trabajo.
- Medidas fundamentales de prestaciones de un servidor
 - Tiempo de respuesta o latencia (R):
 - Tiempo total desde que se solicita una tarea al servidor o a un componente del mismo y la finalización de la misma.
 - Tiempo de ejecución de un programa.
 - Tiempo de lectura de un determinado fichero de un disco.
 - Productividad o ancho de banda (X):
 - Cantidad de trabajo realizado por el servidor o por un componente del mismo por unidad de tiempo.



¿Qué afecta a las prestaciones? ¿Cómo podemos mejorarlas?

Componentes hardware del sistema:

- Características y configuración (cuello de botella).

Sistema operativo:

- Tipo de sistema operativo.
- Políticas de planificación de procesos.
- Configuración de memoria virtual, etc.

Aplicaciones:

- Hot spots.
- Acceso a E/S, swapping, etc.
- Fallos de caché, de página, etc.

Actualización de componentes:

- Reemplazar por dispositivos más rápidos.
- Añadir nuevos componentes: distribución de carga (load balancing). Mayor carga a componentes más rápidos.

Ajuste o sintonización:

- Configuración de componentes hardware.
- Parámetros del sistema operativo.
- Optimización de programas.

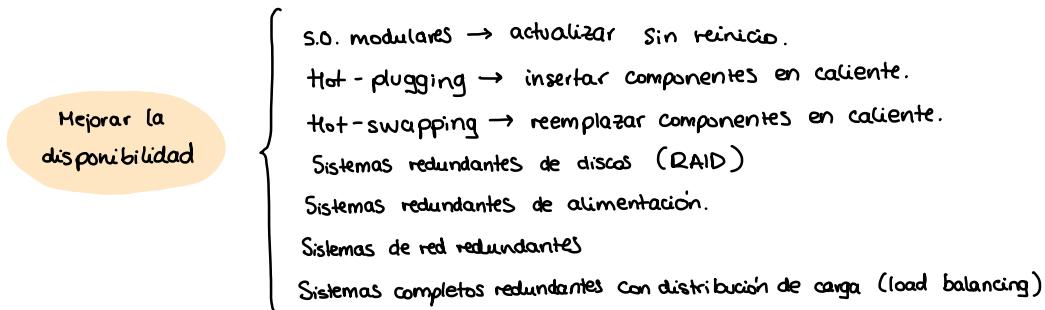
Una de nuestras principales misiones será analizar nuestro servidor para determinar los factores que afectan a su rendimiento y encontrar posibles soluciones para su mejora.

Disponibilidad

Servidor disponible = si se encuentra operativo (respondiendo a peticiones).

Tiempo de inactividad = tiempo en que sistema no disponible.

- Planificado (actualizaciones SW o HW que requieran re-arranques).
- No planificado. Surgen de algún evento físico (fallos en HW, anomalías ambientales, fallos SW)
↳ Alta disponibilidad ⇒ tolerante a fallos.



Fiabilidad

Un sistema es fiable cuando desarrolla su actividad sin errores.

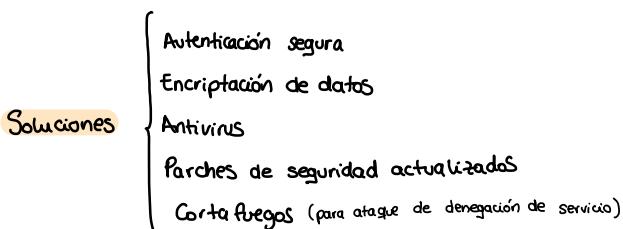
MTTF (Mean time to failure) → tiempo medio que tiene un sistema hasta que ocurre un error.

Soluciones → uso de sumas de comprobación para detección y/o corrección de errores comprobación de recepción.

Seguridad

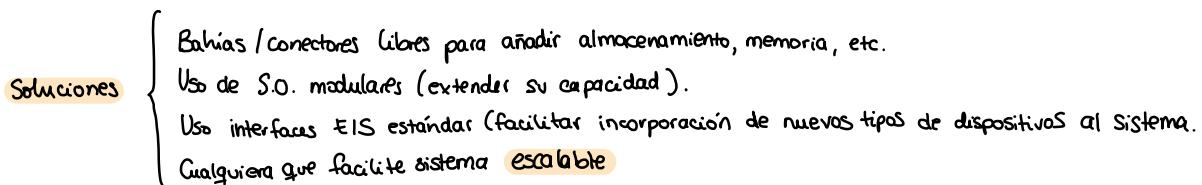
Un servidor debe ser seguro ante:

- Incusión individuos no autorizados → Confidencialidad
 - Corrupción o alteración no autorizada de datos → Integridad
 - Interferencias (Ataques) que impidan el acceso a los recursos.
- } Solución. Encriptación



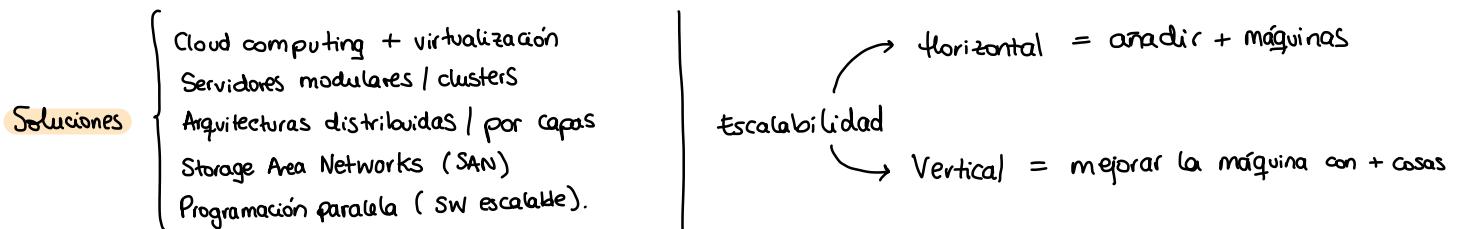
Extensibilidad - Expansibilidad

Facilidad del sistema para aumentar sus recursos.



Escalabilidad

fácilidad para aumentar significativamente los recursos para enfrentarse a un aumento significativo de carga.



! Todos los sistemas ESCALABLES son EXTENSIBLES pero NO a la INVERSA.

Mantenimiento

Acciones con objetivo de prolongar el funcionamiento correcto del sistema

Importante servidor fácil de mantener. Para ello:

- S.O. actualizaciones automáticas
- Cloud computing
- Automatización de copias de seguridad
- Automatización de tareas de configuración y administración

Coste

Adaptarnos al presupuesto. Hay que tener en cuenta el coste HW y SW además de mantenimiento, personal, proveedores de red, alquiler del local donde se ubica el servidor, consumo eléctrico.

Para abaratar el coste:

- Cloud Computing
- SW código libre
- Reducir costes de electricidad (eficiencia energética):
 - Ajuste automático del consumo de potencia de los componentes electrónicos según la carga.
 - free cooling: Utilización de bajas temperaturas exteriores para refrigeración gratuita.

Introducción a la comparación de características entre servidores

Comparación de prestaciones

Computador con mejores prestaciones para un conjunto de programas = aquel que ejecuta dicho conjunto en el tiempo + corto.

Comparar prestaciones → se suele indicar (en lugar de tiempo)

¿Cuántas veces + rápido es uno que otro?

¿Qué % de mejora aporta el más rápido con respecto al otro?

Sea t_A = tiempo ejecución de máquina A (ídem para t_B).

↳ t_R es t_A/t_B veces t_B .

Cambio relativo de t_A con respecto a t_B : $\Delta t_{AB} (\%)$ viene dado por $t_A = t_B + \frac{\Delta t_{AB} (\%)}{100} \cdot t_B$

$$\Delta t_{AB} (\%) = \frac{t_A - t_B}{t_B} \times 100 = \left(\frac{t_A}{t_B} - 1 \right) \cdot 100$$

¿Qué máquina es + rápida? Ganancia en velocidad (speed up)

$$S_B(A) = \frac{v_A}{v_B} = \frac{t_B}{t_A} \rightarrow t_B \text{ llamado tiempo mejorado.}$$

↪ ganancia de A con respecto de B.

$$\Delta v_{AB} (\%) = \left(\frac{v_A}{v_B} - 1 \right) \cdot 100 = (S_B(A) - 1) \cdot 100$$

↓ dominio $[0, +\infty)$

Ejemplo: $t_A = 10s$; $t_B = 5s \Rightarrow t_A / t_B = 2$ veces t_B . t_B es $5/10 = 0.5$ veces t_A .

$$\Delta t_{AB} (\%) = \left(\frac{10}{5} - 1 \right) \cdot 100 = 100\% \Rightarrow t_A \text{ es } 100\% \text{ mayor que } t_B \Rightarrow t_A \text{ es } 2 \text{ veces mayor que } t_B$$

$$\Delta t_{BA} (\%) = \left(\frac{5}{10} - 1 \right) \cdot 100 = -50\% = t_B \text{ es un } 50\% \text{ menor que } t_A.$$

$$\begin{aligned} S_A(B) &= \frac{10}{5} = 2 \\ \Delta v_{AB} (\%) &= (2 - 1) \cdot 100 = 100\% \\ S_B(A) &= \frac{5}{10} = 0.5 \end{aligned} \quad \left. \begin{array}{l} B \text{ es } 2 \text{ veces más rápida que } A \\ B \text{ es un } 100\% \text{ más rápida que } A \\ A \text{ es un } 50\% \text{ más lenta que } B \end{array} \right\}$$

Coste y relación prestaciones/coste

$$\text{Prestaciones/Coste (A)} = \frac{v_A}{\text{Coste}_A} \propto \frac{1/t_A}{\text{Coste}_A} \quad (\text{idem para } B)$$

↓ proporcional

$$\left. \begin{array}{l} t_A = 36s \\ \text{Coste}_A = 625\text{€} \end{array} \right\} \text{Prestaciones/Coste (A)} = \frac{1/36s}{625} = 4.4 \cdot 10^{-5} \text{ s}^{-1}/\text{€}$$

$$\left. \begin{array}{l} t_B = 45s \\ \text{Coste}_B = 550\text{€} \end{array} \right\} \text{Prestaciones/Coste (B)} = \frac{1/45}{550} = 4.0 \cdot 10^{-5} \text{ s}^{-1}/\text{€}$$

A mejor relación prestaciones/coste que B → $\frac{\text{Prestaciones/Coste (A)}}{\text{Prestaciones/Coste (B)}} = 1.1 \Rightarrow A \text{ 1.1 veces mejor que } B$

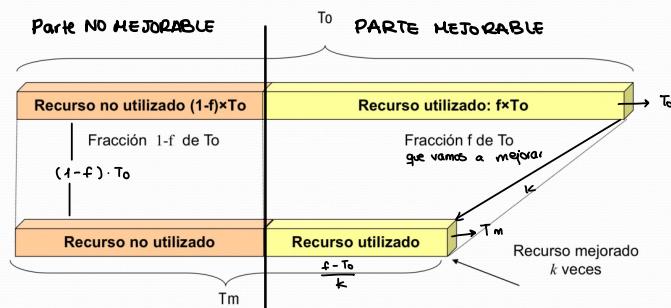
Introducción a la mejora de tiempo de respuesta de un servidor

Mejorar t. respuesta de un servidor = reemplazar un componente del mismo por otro + rápido.

Suponemos que servidor sólo ejecuta un proceso monohebra. En ese caso:

- Servidor tarda un tiempo $T_{original} = T_0$ en ejecutar el proceso
- Mejoramos el sistema reemplazando uno de los componentes por otro k veces "+ rápido" (suponiendo $k > 1$) es decir un $(k-1) \cdot 100$ + rápido.
- $T_f = \text{Tiempo mejora}$
- Este componente se utiliza durante una fracción f de T_0 ($f=1 = 100\%$ del tiempo)

¿Cuál es la ganancia que conseguimos en la ejecución de ese proceso?



LEY DE AMDAHLL

$$S = S_0(m) = \frac{v_m}{v_0} = \frac{T_0}{T_m} = \frac{T_0}{(1-f) \cdot T_0 + f \cdot \frac{T_0}{k}}$$

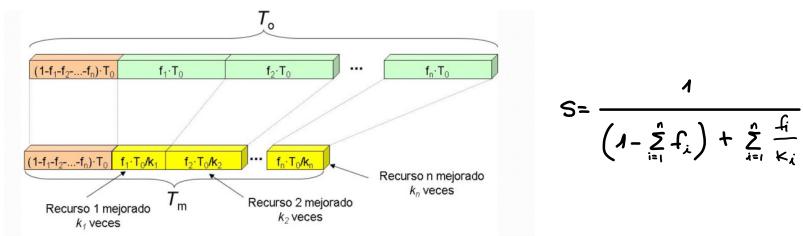
$$S = \frac{1}{1-f + f/k}$$

S; $f=0 \Rightarrow S=1$ (no hay mejora)
S; $f=1 \Rightarrow S=k$ (mejora tantas veces como el recurso utilizado)

$$S_{max} = \lim_{k \rightarrow \infty} S = \frac{1}{1-f}$$

WUOLAH

Para n mejoras:



CONCLUSIONES

mejora efectiva cuanto > es f . \Rightarrow hay que optimizar los elementos durante la mayor parte del tiempo.