

## Clase 19/10/2020

### Cuando se requiere sincronizar

El tiempo es una referencia que utilizan los sistemas distribuidos en varias situaciones.

Supongamos una plataforma de comercio electrónico que funciona a nivel global, en cada país se tiene un servidor con una base de datos dónde se registran las compras, incluyendo la fecha y hora en la que se realiza cada compra.

Para consolidar las compras a nivel mundial cada servidor debe enviar los datos a un servidor central. Sin embargo, no es posible ordenar las compras por fecha debido a dos situaciones:

1. Cada compra se ha registrado con la fecha y hora local, y
2. No es posible garantizar que los relojes de los servidores funcionen a la misma velocidad.

Para ilustrar este problema supongamos que un cliente en México realiza una compra a las 8 PM, y un cliente en España realiza una compra a las 2 AM del día siguiente ¿quién compró primero?

Aparentemente el cliente en México realizó la compra antes que el cliente en España, debido a que la fecha de la compra del cliente en México es un día anterior a la fecha de la compra del cliente en España. Sin embargo, en realidad el cliente en España realizó la compra una hora antes que el cliente en México, debido a que la diferencia horaria entre México y España es de 7 horas.

### Sincronización de relojes

Sincronizar dos o más relojes significa que los servidores se ponen de acuerdo en una misma hora. Notar que un grupo de servidores pueden ponerse de acuerdo en una hora y otro grupo de servidores puede ponerse de acuerdo en otra hora; solo si ambos grupos de servidores se conectan entonces ambos grupos deberán acordar una hora.

Como se dijo anteriormente, **el tiempo es una referencia para establecer un orden** en una secuencia de eventos (como serían las compras en una plataforma de comercio electrónico).

Más adelante veremos que éste orden puede establecerse utilizando relojes físicos (mecanismos que marcan el tiempo real) o bien relojes lógicos (contadores).

## Relojes físicos

En los sistemas digitales, un reloj físico es un circuito que genera pulsos con un periodo "constante".

En una computadora cada pulso de reloj produce una interrupción en el CPU para que se actualice un contador de "ticks". Dado que el pulso tiene un periodo "constante" el número de ticks es una medida del tiempo transcurrido desde que se encendió la computadora.

El siguiente diagrama muestra un reloj físico el cual genera pulsos regulares. Cuando la señal cambia de 0 volts a 5 volts se produce una interrupción en el CPU, entonces se invoca una rutina llamada manejador de interrupción (*handler*) la cual incrementa el contador de "ticks".

El contador de "ticks" de una computadora no es un reloj preciso, dado que:

1. Los relojes físicos se construyen utilizando cristales de cuarzo con la finalidad de tener un periodo de oscilación constante, sin embargo los cambios en la temperatura modifican el periodo del pulso, lo que ocasiona que el reloj se adelante o se atrase.
2. Cuando se produce la interrupción al CPU, el sistema podría estar ejecutando una rutina de mayor prioridad, por tanto la rutina que incrementa los "ticks" se bloquea lo que provoca que algunos pulsos de reloj no incrementen la cuenta de "ticks".

## Segundos solares

El concepto de tiempo que utilizamos en la práctica se basa en la percepción que tenemos del día. Un día es un período de luz y oscuridad debido a la rotación de la tierra sobre su eje.

Dividimos convencionalmente el día en 24 horas, cada hora en 60 minutos y cada minuto en 60 segundos. Por tanto, la tierra tarda 86,400 segundos en dar una vuelta sobre su eje, en términos de velocidad angular estamos hablando de  $360/86400=0.00416$  grados/segundo. Así, a la fracción  $1/86400$  de día le llamamos **segundo solar**.

## Segundos atómicos

Una forma más precisa de medir el tiempo es utilizar un reloj atómico de Cesio 133.

En un reloj atómico se aplica microondas con diferentes frecuencias a átomos de

Cesio 133, entonces los electrones del átomo de Cesio 133 absorben energía y cambian de estado; posteriormente los átomos regresan a su estado basal emitiendo fotones.

A la frecuencia que produce más cambios de estado en los electrones del átomo de Cesio 133 se le llama *frecuencia natural de resonancia*.

La frecuencia natural de resonancia del Cesio 133 es de 9,192,631,770 ciclos/segundo, es decir, el átomo de Cesio 133 muestra un máximo de absorción de energía cuando se le aplica microondas con una frecuencia de 9,192,631,770 Hertzs.

Entonces se define el **segundo atómico** como el recíproco de la frecuencia natural de resonancia del Cesio 133 (recordar que el periodo de una onda es el recíproco de su frecuencia).

Los relojes atómicos de Cesio 133 son extremadamente precisos, ya que independientemente de las condiciones ambientales (temperatura, presión, etc.), se adelantan o atrasan un segundo cada 300 millones de años.

## **Tiempo atómico internacional TAI**

Se define el tiempo atómico internacional (TAI) como el promedio de los segundos atómicos transcurridos desde el 1 de enero de 1958, dicho promedio obtenido de casi 70 relojes de Cesio 133 al rededor del mundo.

## **Tiempo universal coordinado UTC**

El tiempo universal coordinado UTC (*Coordinated Universal Time*, CUT) es el estándar de tiempo que regula actualmente el tiempo de los relojes a nivel internacional.

El tiempo UTC ha reemplazado el tiempo tiempo medio de Greenwich GMT.

El tiempo GMT toma como referencia la posición del sol a medio día. Tanto el tiempo GMT como el tiempo UTC consideran el día solar compuesto por 86400 segundos solares.

Debido a que nuestro planeta disminuye su velocidad angular lentamente, el segundo solar dura más que el segundo atómico. Para sincronizar los segundos UTC con los segundos TAI, el tiempo UTC se debe "atrasar" con respecto al tiempo TAI, para esto "se salta" un segundo UTC una vez al año; se dice entonces que se introducen **segundos vacíos** en el tiempo UTC.

Los proveedores de nube han adoptado el uso del tiempo UTC para los relojes en las máquinas virtuales, por ejemplo cuando se ejecuta el comando **date** en una máquina virtual con Ubuntu en Azure, se obtiene la fecha y hora UTC.

## Sincronización de relojes físicos

En un sistema centralizado el tiempo se obtiene del reloj central, por tanto todos los procesos se sincronizan mediante un sólo reloj.

En un sistema distribuido cada nodo tiene un reloj que se atrasa o adelanta dependiendo de diversos factores físicos. A la diferencia en los valores de tiempo de un conjunto de computadoras se le llama **distorsión del reloj**.

## Network Time Protocol NTP

El protocolo de tiempo de red (*Network Time Protocol, NTP*) define un procedimiento centralizado para la sincronización de relojes. En este procedimiento los clientes consultan un servidor de tiempo, el cual podría contar con un reloj atómico o estar sincronizado con una computadora que tenga un reloj atómico.

El protocolo NTP estima el tiempo que tarda en llegar al servidor de tiempo la petición del cliente  $T_{req}$  y el tiempo que tarda en llegar al cliente la respuesta del servidor  $T_{res}$ .

Supongamos que al tiempo local  $T_1$  el cliente A envía una petición al servidor B, la petición llega al servidor al tiempo local  $T_2$ . El servidor B procesa el requerimiento y al tiempo local  $T_3$  envía la respuesta al cliente A, la respuesta llega al cliente al tiempo local  $T_4$ .

Si el servidor B envía  $T_3$  y  $T_2$  a la computadora A y suponemos  $T_{req}=T_{res}$ , entonces la computadora A puede estimar  $T_{res}$  ya que conoce  $T_4-T_1=T_{req}+T_{res}+(T_3-T_2)=2T_{res}+(T_3-T_2)$ , por tanto la computadora A podrá modificar su tiempo local a  $T_3+T_{res}$ .

Debido a que los relojes atómicos son recursos muy costosos y con el fin de evitar la saturación del servidor que cuenta con un reloj atómico, se suele implementar el mismo procedimiento sobre una topología de árbol.

Los servidores de los estratos superiores del árbol son más exactos que los servidores de estratos inferiores, de tal manera que el servidor en la raíz, llamado

**servidor de estrato 1**, contará con un reloj atómico (llamado reloj de referencia).

Para instalar NTP en Ubuntu, se debe ejecutar los siguientes comandos:

```
sudo apt-get update  
sudo apt-get install ntp
```

### **Algoritmo de sincronización de relojes de Berkeley**

En el algoritmo NTP el servidor es pasivo, ya que espera recibir las peticiones de los cliente.

El algoritmo de sincronización de relojes de Berkeley es un procedimiento descentralizado donde el servidor tiene una función activa, ya que cada cierto tiempo inicia la sincronización de un grupo de computadoras.

El algoritmo de Berkeley se basa en el principio que enunciamos al principio: *si dos computadoras no están conectadas, entonces no requieren sincronizar sus tiempos*.

Por tanto, si un grupo de computadoras no se conectan con otras computadoras, es suficiente sincronizar los tiempos de las computadoras en el grupo, aún si el tiempo sincronizado no corresponde al tiempo real (ya que no hay una comunicación con otras computadoras).

En la práctica no hay computadoras aisladas del mundo real, de manera que el algoritmo de Berkeley se puede utilizar para sincronizar las computadoras de una red local, mientras que alguna de las computadoras se podría sincronizar con un servidor de tiempo utilizando NTP.

El algoritmo de Berkeley es el siguiente:

1. El nodo A (servidor) le envía a los nodos A, B y C su tiempo.
2. Los nodos A, B y C les envían al nodo A las diferencias entre sus tiempos y el tiempo en el nodo A.
3. El nodo A calcula el promedio de las diferencias. El nodo A envía a los nodos A, B y C la corrección de tiempo.
4. Los nodos A, B y C modifican sus tiempos locales.