

Coherencia de caché - Paquete 1

Francisco Gil Amorós

UCA

- 1 **Introducción**
- 2 Estrategias de escritura. WT y WB
- 3 Coherencia en monoprocesadores. Problemas con DMA
- 4 Bibliografía

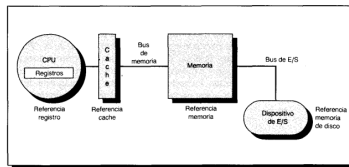
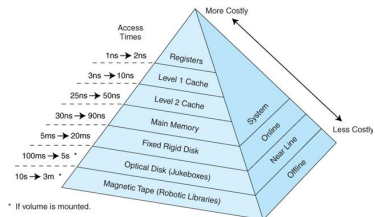
Caché

Caché : es un sitio seguro para ocultar o almacenar cosas. *Nuevo Diccionario Mundial de Webster del Lenguaje Americano. Segunda edición de Colegio (1976)*

En informática, una **caché** es un componente (HW/SW) que **almacena datos** para que las solicitudes futuras de esos datos se puedan atender con **mayor rapidez**. Los datos almacenados pueden ser el resultado de un **cálculo anterior** o el **duplicado de datos** almacenados en otro lugar, generalmente, da velocidad de acceso más rápido.

Caché nombre inicialmente escogido para representar el nivel de **jerarquía de memoria** entre la **CPU** y **memoria principal**.

Jerarquía de memoria



Memorias : - **capacidad** \Rightarrow + **rápidas y caras**

Una **jerarquía de memoria** consta de muchos **niveles** y en cada momento se gestiona entre **dos niveles adyacentes**. Cuanto más nos acercamos al procesador (*nivel superior*) usaremos memorias **más rápidas** como la caché.

Bloque : **Mínima unidad de información** que puede estar presente o no presente en la jerarquía de dos niveles. Tamaño **fijo** o **variable**.

¿Qué es y por qué?

La caché contiene **copia de datos** en memoria, nunca datos propios. Son bloques con altas probabilidades de que se usen.

Se produce un **acierto de caché** cuando el procesador intenta acceder a un dato (siempre está en memoria, a veces replicado en memoria caché). Si está en caché ésta es la que responde al procesador.

¿Qué es y por qué?

¡**PROBLEMA!** Tener **copias** nos hace **difícil** asegurar que todas las copias están **actualizadas**.

Con la **coherencia de caché** se garantiza que las copias están **actualizadas** y -obviamente- son **iguales**.

Conceptos y términos usados

¿Sistema de memoria es **coherente**?

Operación de lectura retorna **siempre** el **último valor independientemente** del **procesador** que efectúa lectura o escritura.

Si un dato se encuentra en el nivel más alto de la jerarquía y ha sido modificado, activará **mecanismos** para que esta **modificación** se lleve a cabo en el **resto de niveles**, donde está también ese dato.

- 1 Introducción
- 2 Estrategias de escritura. WT y WB**
- 3 Coherencia en monoprocesadores. Problemas con DMA
- 4 Bibliografía

Estrategias de escritura

Las **estrategias de escritura** establecen la relación entre las escrituras en la memoria caché (Mc) y la memoria principal (Mp).

Objetivo : garantizar coherencia entre información de la Mc y la Mp.
Rendimiento del sistema se vea afectado lo mínimo posible.

Dos estrategias \implies **WT** y **WB**

Write through

Write through (WT) o escritura inmediata

- **Segura.** Al escribir un dato en Mc, se **modificará inmediatamente** en Mp. **Elimina problemas** de coherencia, **evita** tener que **comprobar** si el dato está cacheado y si es diferente al que hay en Mp.
- **Poco eficiente.** No hay ganancia de tiempo en las escrituras (sólo lecturas), ya que cada escritura en caché en realidad escribe en Mp (+ costoso).

Write back

Write back (WB) o escritura diferida/retardada

- **Más eficiente.** Datos inicialmente se escriben en caché. Se **actualizan** en **Mp sólo cuando** el bloque que lo contiene es **reemplazado**, es decir, cuando hay que eliminar el bloque de la Mc.
- **Complejo mantener segura.** Aparecen **problemas de coherencia**, no siempre es coherente si no es necesario.

Se asocian **dos bits** más a cada bloque de caché :

- **bit dirty** cuando la CPU escribe en el bloque de caché (dato modificado/actualizado).
- **bit valid** para comprobar si el dato contiene información válida/útil.

WT y WB

Write back proporciona un **+** **rendimiento** que la Write through, a costa de una **+** **complejidad** HW para mantener coherencia.

Con **Write back** realizamos - **accesos a memoria**. Se suele usar WB en **monoprocesadores**.

- 1 Introducción
- 2 Estrategias de escritura. WT y WB
- 3 Coherencia en monoprocesadores. Problemas con DMA**
- 4 Bibliografía

Coherencia en monoprocesadores. Problemas con DMA

La **DMA** puede llevar a **problemas de coherencia de caché**.

Caso : CPU con memoria caché y una memoria externa que pueden acceder directamente los dispositivos (DMA).

Cuando la CPU accede a X lugar en la memoria, el valor actual se almacena en la caché. Si se realizan operaciones posteriores en X, se **actualizará la copia en caché** de X, **pero no la versión de memoria externa** de X.

Si la **caché no se vacía** en la memoria antes de que otro dispositivo intente acceder a X, el dispositivo recibirá un valor **caducado** de X.

- 1 Introducción
- 2 Estrategias de escritura. WT y WB
- 3 Coherencia en monoprocesadores. Problemas con DMA
- 4 Bibliografía**

Bibliografía



Arquitectura de Computadores - Un enfoque cuantitativo
David A. Patterson, John L. Hennessy.



Coherencia de caché.

Apuntes Universidad Europea de Madrid. http://www.cartagena99.com/recursos/alumnos/apuntes/ININF1_M10_U4_T3.pdf



Multiprocesadores : Coherencia, consistencia y sincronización

Apuntes Universidad Complutense de Madrid. https://www.fdi.ucm.es/profesor/rhermida/AC-Grado/AC_tema6.pdf



Acceso directo a memoria - Problemas de coherencia en la memoria caché
Wikipedia. <https://bit.ly/2YBAfJS>