

# Memòria caché (2)

## Problema

La informació que tenim a les caches és un “duplicat” d’informació que està a MP. Si aquesta informació és compartida amb altres dispositius hem de tenir cura que sempre es faci servir la “còpia” correcta.

On poden sorgir problemes?:

- Sistemes multiprocessador o multi-core (cada un amb la seva cache).
- Sistemes que encara que només tinguin un processador tinguin altres “mestres” capaços d’accedir/modificar la informació independentment del processador (CDMA, co-processadors...)

Si una dada està a més d’un lloc al sistema i algun element modifica la dada a un dels llocs on es troba (MP, o alguna de les caches) tindrem la mateixa informació amb diferents valors, llavors:

- El valor correcte sempre serà l’últim que s’ha actualitzat.
- Si algun altre dispositiu ha de fer servir aquesta informació, el sistema ha de subministrar l’última actualització.

## Solucions

A aquests tipus de sistemes direm que, per no tenir problemes, hem de mantenir **Coherència de la Informació** en tot el sistema de memòria (MP+Cachés).

Hi ha diversos tipus de solucions, però bàsicament consisteixen en que es passi informació dels canvis entre Memòria Principal i les Cachés.

Aquest sistema de notificació i atenció entre memòries es denomina **Bus Snooping**.

Degut a que la MP és passiva, seran els sistemes de caché amb els seus controladors els que hauran d'enviar els missatges de canvis. A la vegada, han “d'escoltar” dels busos (*snoop*) els missatges enviats per les altres.

## Protocol amb Política d'Espectura *Write-Through*

Aquest protocol és molt **senzill** i tant sols necessita el bit de estatus: “*Valid bit*”.

Totes les escriptures a caché es fan també a MP:

- A M.P. sempre tindrem la informació actualitzada.
- Les cachés que “escoltin” (*snoop*) que una informació que elles tenen ha estat modificada per algú altre, posen el “*Valid bit*”=“invàlid”. Si s'intenta llegir aquesta informació aquest bit generarà un MISS.

TAULA D'ACTIVITAT de les CACHÉS

Operació a la Caché		Activitat al Bus	Les altres Cachés	Comentaris
Lectura	Hit Lectura	cap	No passa res	La dada és llegeix de caché
Lectura	Miss Lectura	Lectura d'un bloc	No passa res	La dada és llegeix de MP i s'actualitza a caché
Espectura	Hit Espectura	Espectura a MP	Snooping→ <i>Valid_bit</i> =Invàlid	La dada s'escriu a caché i a MP. Les altres cachés invaliden la seva dada
Espectura	Miss Espectura	Espectura a MP	Snooping→ <i>Valid_bit</i> =Invàlid	La dada s'escriu a MP. Les altres cachés invaliden la seva dada

**Punts Febles: Lent, genera molt moviment de dades al bus del sistema.**

## Protocols amb Política d'Espectura *Write-Back*

El protocol anterior és molt senzill però té el gran inconvenient que genera molt de moviment de dades al bus del sistema.

Per evitar aquesta situació necessitem protocols de coherència que puguin treballar amb la política d'espectura *Write-Back*.

La idea serà mantenir una línia de caché, encara que es modifiqui la seva informació respecte a la MP, mentre cap altre element del sistema necessiti aquesta informació (és a dir mentre sigui informació exclusiva).

Només s'actualitzarà la informació, a la MP i a altres caches, quan la necessitin.

S'han d'afegir línies de control addicionals per informar d'accions com:

- Lectura
- Espectura/Modificació
- Tinc còpia d'una informació (bona o dolenta)...

Hi ha diversos protocols d'aquest tipus, però el més utilitzat és el MESI.

## Protocol MESI

- Mentre una dada sigui “propietat” d’una caché no es transmetran fora (*broadcasted*) les modificacions. Invalidarà còpies a altres cachés.
- Les modificacions de dades que no siguin “propietat”, és a dir que són compartides, si seran “*broadcasted*”.

Cada línia de caché té assignat un de quatre possibles estats:

- **Modificat (M):** La dada és bona. Tant sols està a aquesta caché, a més, ha estat modificada respecte a M.P. (Write-Back)
- **Exclusiu (E):** La dada és bona. Tant sols està a aquesta caché, i coincideix amb el contingut de M.P. (Write-Back)
- **Compartit (S):** La dada és bona, però ara pot estar a altres cachés. Sempre conté la dada actualitzada. (Write-Through).
- **Invalit (I):** La dada és incorrecte, s’ha d’anar a llegir a la M.P. Genera un caché miss.

Protocol MESI

Taula Resum:

	<b>M</b> Modificat	<b>E</b> Exclusiu	<b>S</b> Compartit	<b>I</b> Invàlid
<b>És vàlida la informació a la línia de caché?</b>	Si	Si	Si	No
<b>La dada a MP és :</b>	No vàlida	Vàlida	Vàlida	Depèn d'altres cachés
<b>Hi ha còpies d'aquesta dada a altres cachés?</b>	No	No	Potser	Potser
<b>Si escrivim a aquesta línia:</b>	Només a la caché. No surt al Bus	Només a la caché. No surt al Bus	S'actualitza la caché. A més, surt al Bus.	Surt al bus, per actualitzar el subsistema de memòria

Protocol MESI

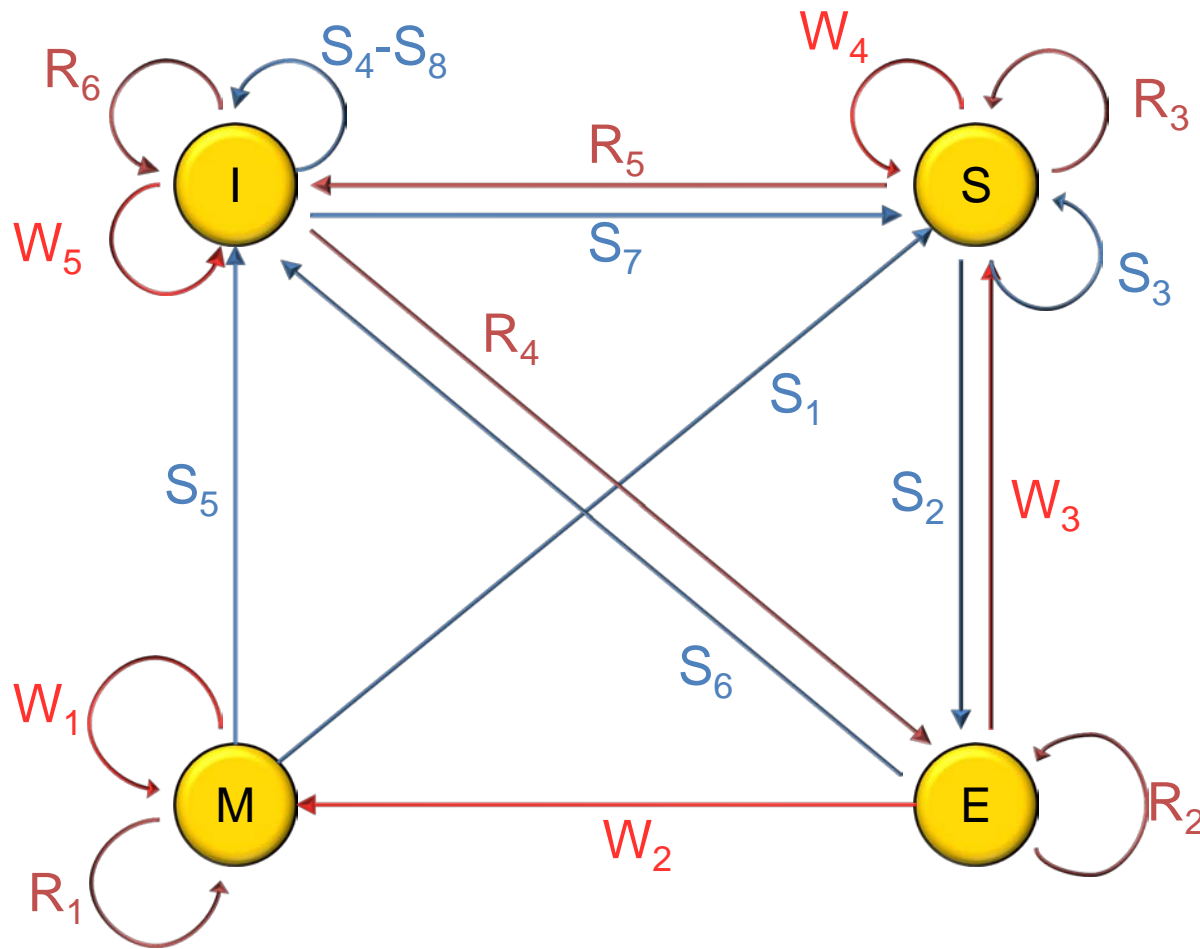
Possibles transicions durant les diferents operacions:

Lectures		Escriptures		Snooping Lectura	
<b>M→M</b> (R <sub>1</sub> )	Dada correcte a caché Es manté al mateix estat	<b>M→M</b> (W <sub>1</sub> )	Write-Back	<b>M→S</b> (S <sub>1</sub> )	Obliga a actualitzar la M.P. abans de fer l'operació (HitM)
<b>E→E</b> (R <sub>2</sub> )	Dada correcte a caché Es manté al mateix estat	<b>E→M</b> (W <sub>2</sub> )	Write-Back	<b>E→S</b> (S <sub>2</sub> )	Envia tinc aquesta dada a altres cachés (Hit)
<b>S→S</b> (R <sub>3</sub> )	Dada correcte a caché Es manté al mateix estat	<b>S→E</b> (W <sub>3</sub> )	Write-Through Envia invàlid (INV) a altres cachés	<b>S→S</b> (S <sub>3</sub> )	Envia tinc aquesta dada a altres cachés (Hit)
<b>I→E</b> (R <sub>4</sub> )	Es demana la dada a MP o altre caché S'actualitza la caché.			<b>I→I</b> (S <sub>4</sub> )	
<b>I→S</b> (R <sub>5</sub> )	Es demana la dada a MP o altre caché S'actualitza la caché.	<b>S→S</b> (W <sub>4</sub> )	Write-Through Envia invàlid (INV) a altres cachés	Snooping Escriptura	
		<b>I→I</b> (W <sub>5</sub> )	S'escriu directament a M.P. Envia invàlid (INV) a altres cachés	<b>M→I</b> (S <sub>5</sub> )	Obliga a actualitzar la M.P. abans de fer l'operació (HitM)
<b>I→I</b> (R <sub>6</sub> )	Dada a MP. No s'actualitza la caché. Aquesta transició no sempre s'implementa.			<b>E→I</b> (S <sub>6</sub> )	Invalida la seva còpia de la dada
				<b>S→I</b> (S <sub>7</sub> )	Invalida la seva còpia de la dada
				<b>I→I</b> (S <sub>8</sub> )	



## Protocol MESI

Possibles transicions durant les diferents operacions:



R<sub>i</sub> : Lectura.

W<sub>i</sub> : Escriptura.

S<sub>i</sub> : "Snooping".

## Protocol amb *Bus Snooping* amb dos nivells de Caché

Quan fem servir *bus snooping* per fer protocols de coherència amb caches, s'han d'afegir senyals de control entre les caches (hit, hitM, invalidar, preguntar....):

