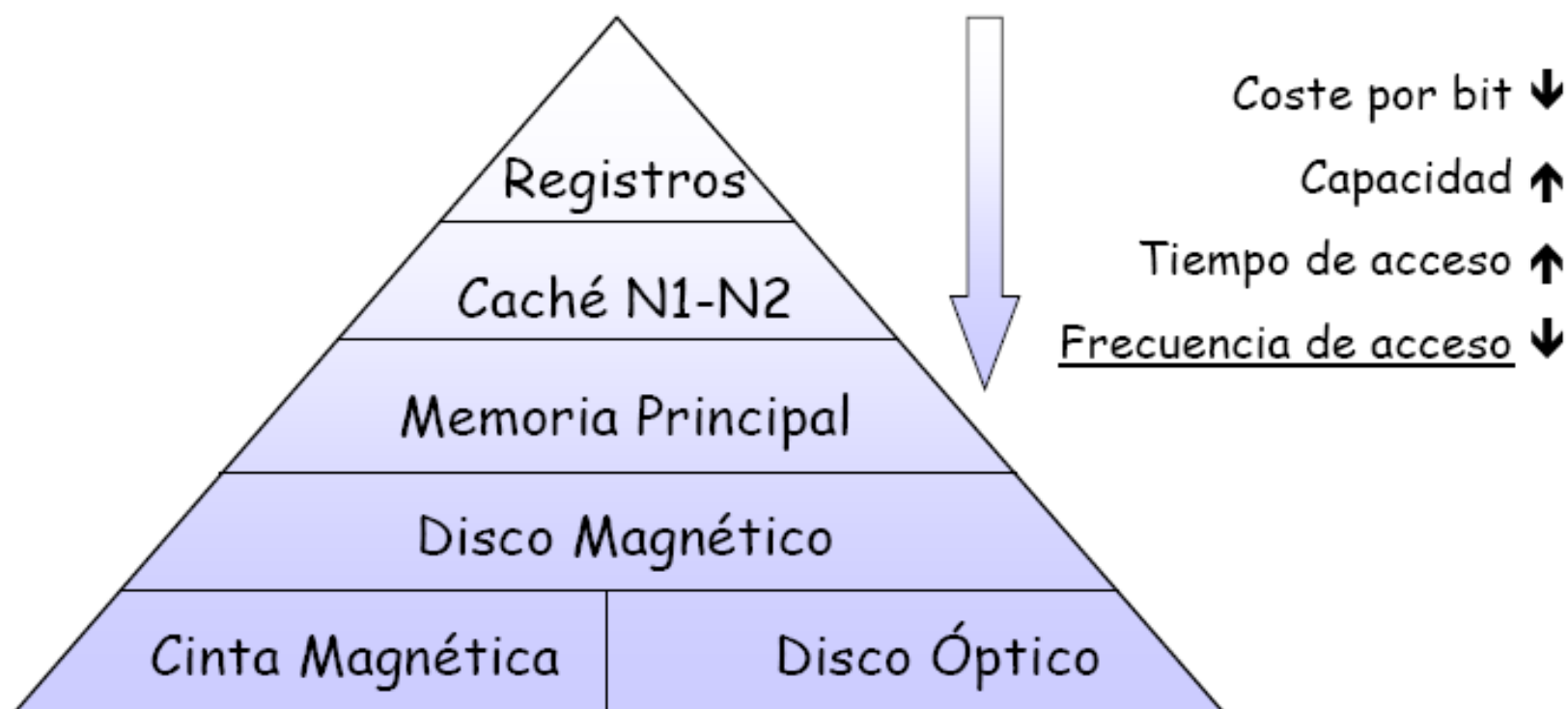


Memoria Caché



AP: Curso 2021/2022



¿Quién se ocupa
del transvase
entre memorias?

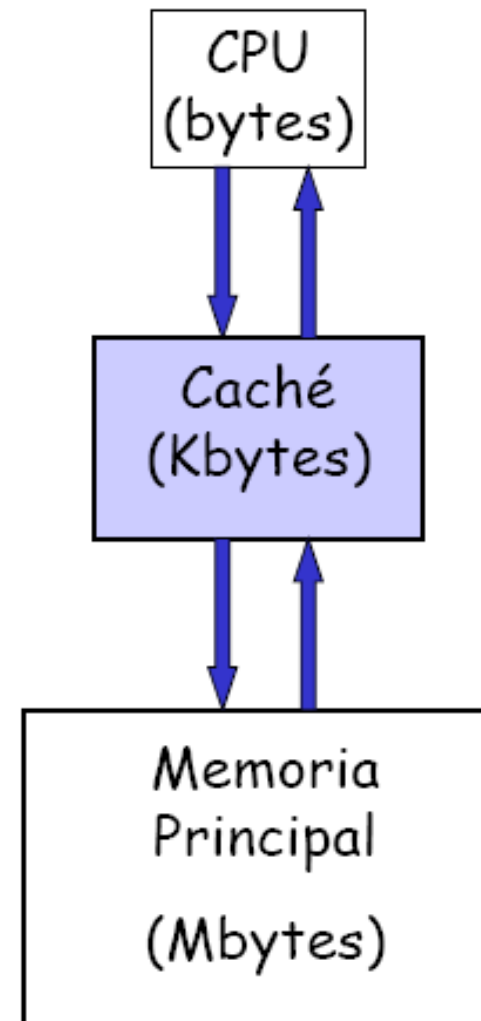


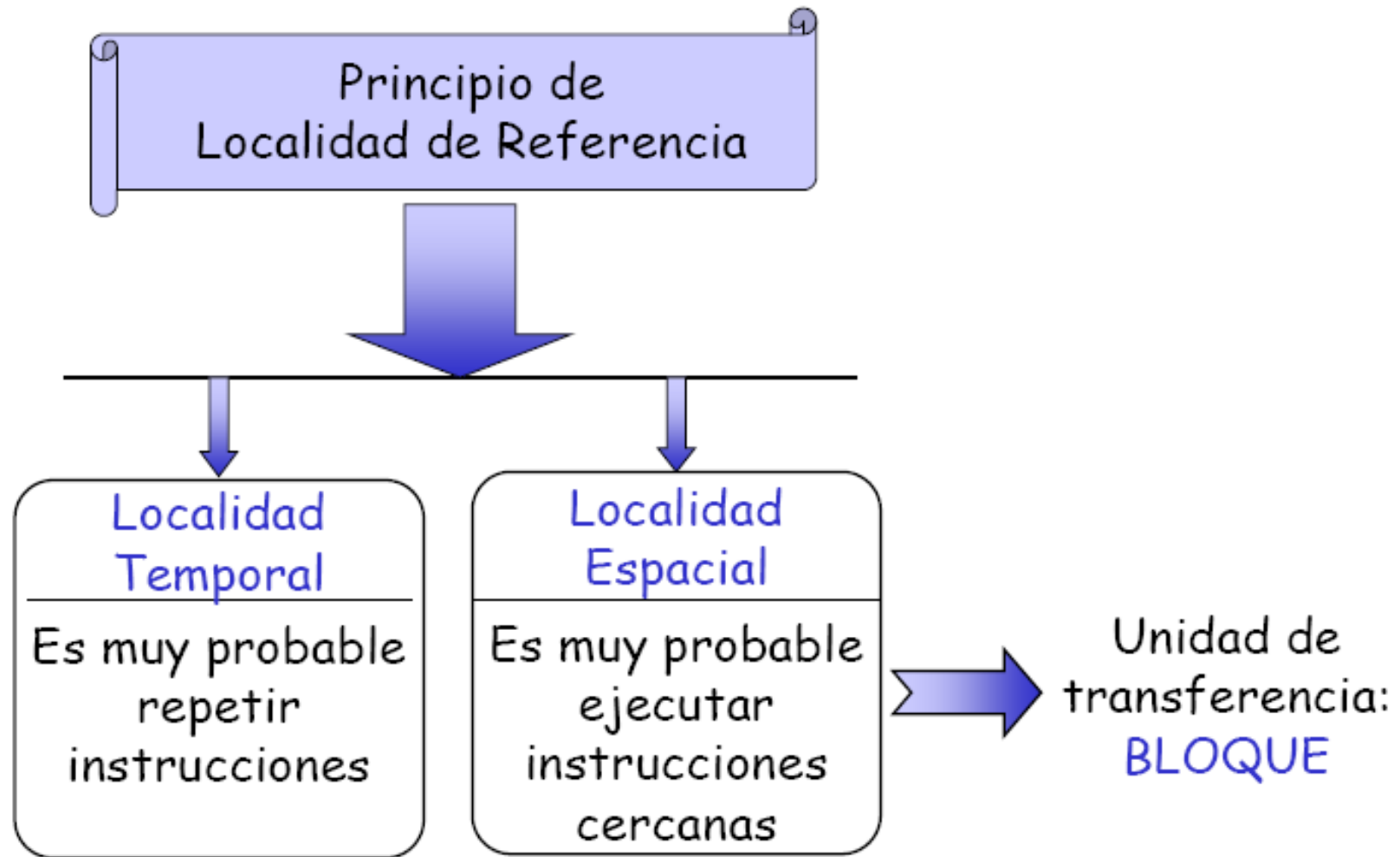
- La CPU/MMU
- El S. O.

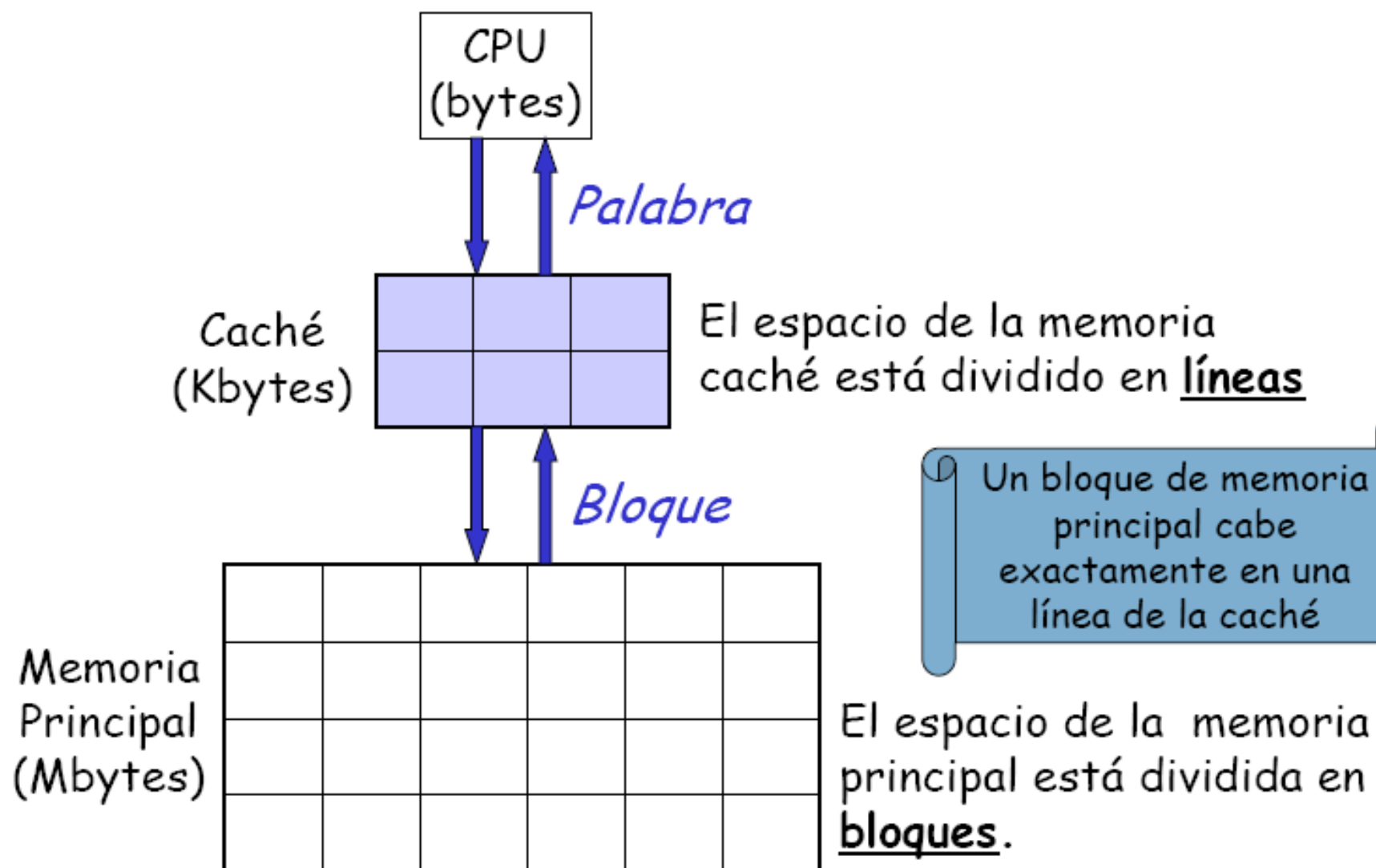
¿ Por qué es tan útil
una memoria tan pequeña?

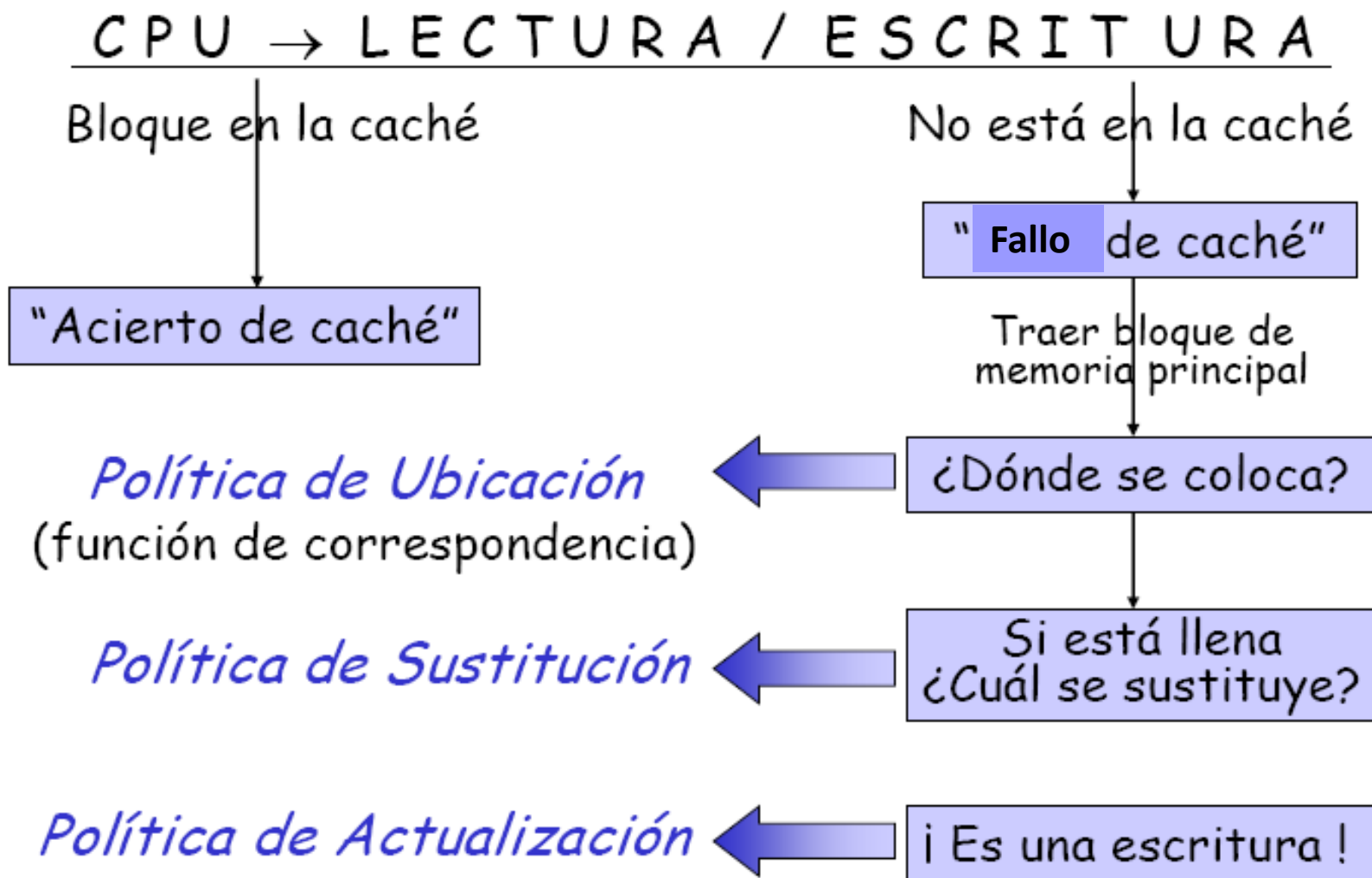
*Principio de la
Localidad de Referencia*

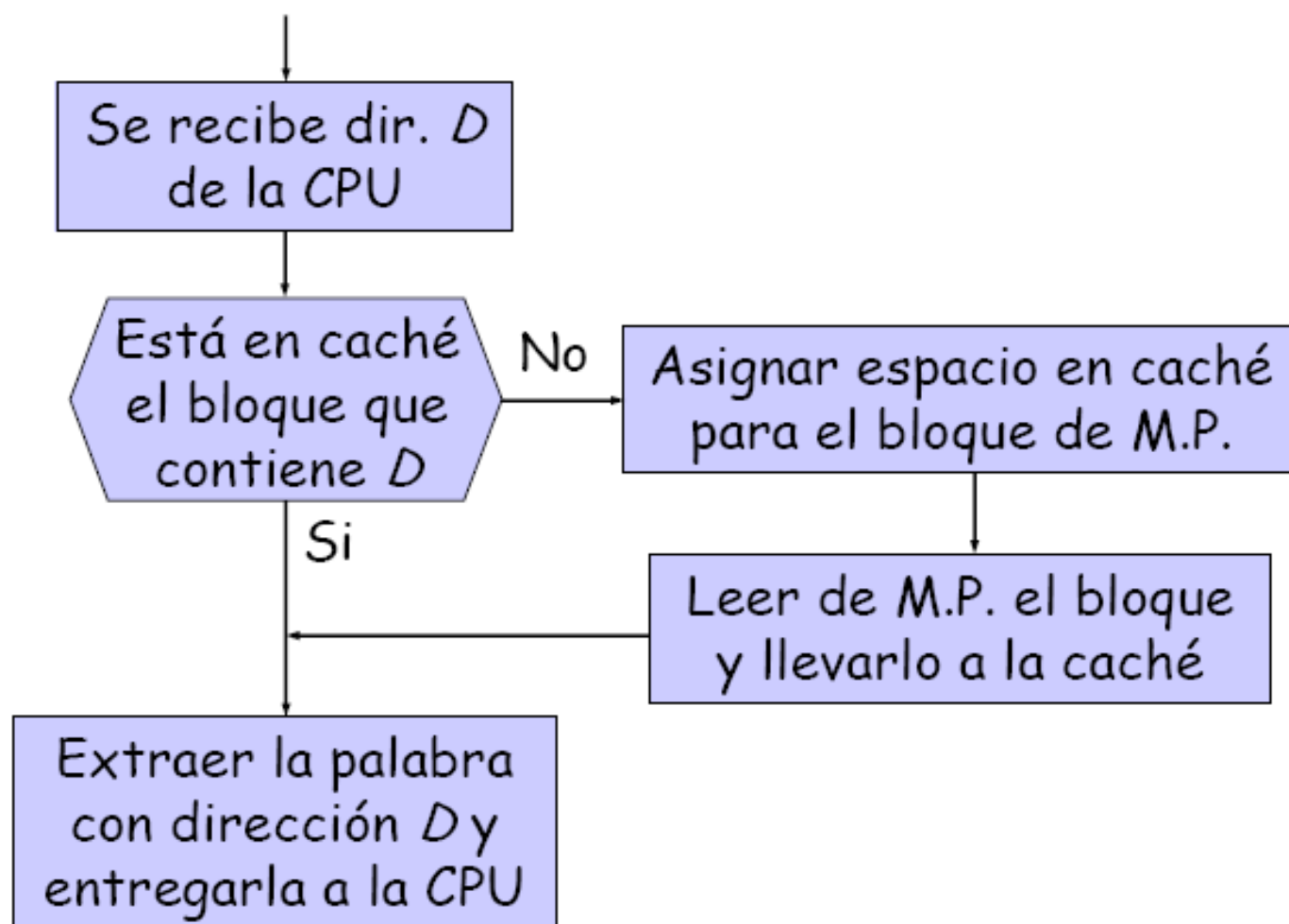
Los accesos a memoria que realiza la CPU no están uniformemente distribuidos por todo el espacio de direccionamiento, sino que se concentran, temporalmente, solamente en ciertas áreas de la memoria.









Operación de lectura en un sistema con caché

Efectividad de la Memoria Caché

Tiempo medio de acceso a memoria

$$T_{\text{acceso}} = T_C \cdot P_A + T_{MP} \cdot (1 - P_a)$$

P_A : Probabilidad de acierto

T_C : Tiempo de acceso a caché

T_{MP} : Tiempo de acceso a M. P.

EJEMPLO

$$T_{MP} = 500 \text{ ns}$$

$$T_C = 50 \text{ ns}$$

$$P_A = 0,99$$

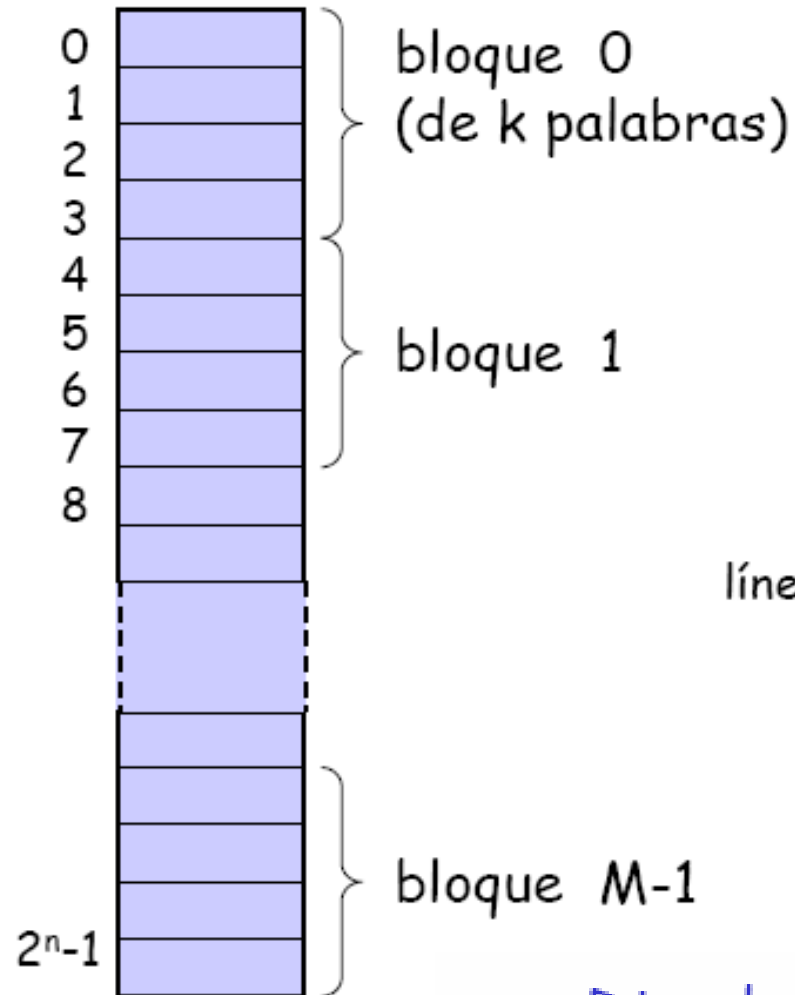
$$T_{\text{acceso}} = 50 \cdot 0,99 + 500 \cdot 0,01 = 54,5 \text{ ns}$$

¿ Merece la pena la caché ?

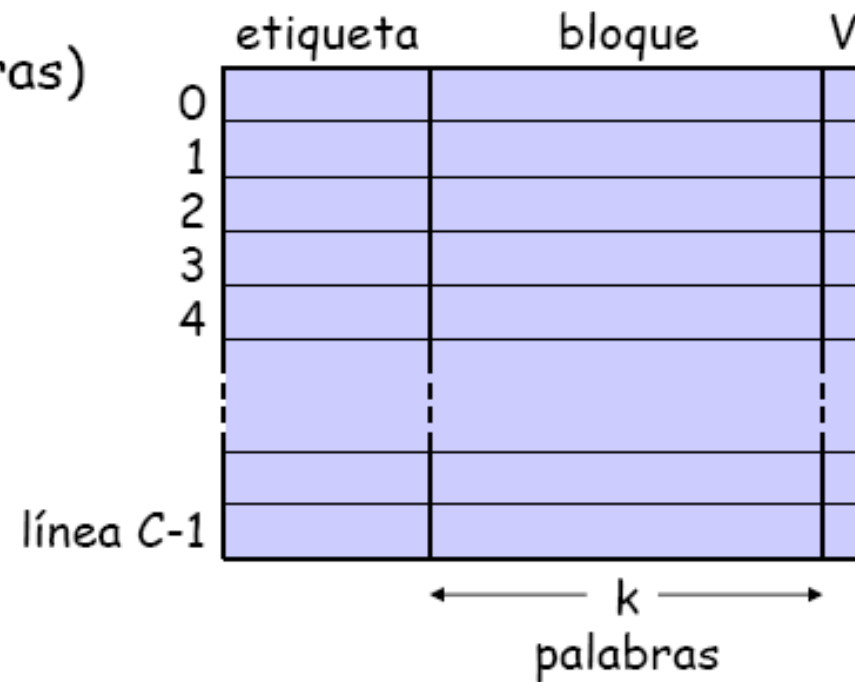
$$\text{Índice de mejora} = \frac{T_{\text{sin caché}}}{T_{\text{con caché}}} = \frac{500}{54,4} = 9,19$$

$T_{MP} = 450 \text{ ns}$ (copia de bloque de MP a MC) + 50 ns (lectura de MC)

Dir. Mem. Principal



Mem. Caché



Bloques >> Líneas

Dir. de memoria

Bloque	Despl.
--------	--------

Datos de los ejemplos

- El tamaño de la caché es de 4 KiBytes
- Los datos se transfieren entre la MP y la MC en bloques de 4 Bytes => 1024 líneas
- La MP es de 64 KiBytes, pudiendo direccionar a nivel de Bytes => dirección de 16 bits, 16 KiBloques de 4 Bytes

Mem. Principal

bloque 0
bloque 1
bloque 2
bloque 3
bloque 4
bloque 5
bloque 6
bloque 7
...
...
...
bl. 16.381
bl. 16.382
bl. 16.383

Caché

etiqueta

línea 0	
línea 1	
línea 2	
línea 3	
...	
...	
línea 1022	
línea 1023	

Función de
correspondencia

¿ En qué línea de la caché
se coloca cada bloque ?

POLÍTICA DE UBICACIÓN
(Función de Correspondencia)

Directa

Asociativa

Asociativa de
Conjuntos

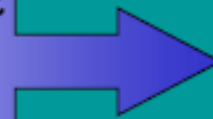
Correspondencia asociativa

- ⌘ Un bloque de memoria principal pueda cargarse en cualquier línea de la cache.
- ⌘ La dirección de memoria se interpreta como una etiqueta y una palabra.
- ⌘ La etiqueta identifica unívocamente un bloque de memoria.
- ⌘ Todas las etiquetas de líneas se examinan para buscar una coincidencia.
- ⌘ La búsqueda de cache es costosa.

Función de Correspondencia

El bloque puede ubicarse en cualquier línea de la caché

Falta de caché
+
Caché llena

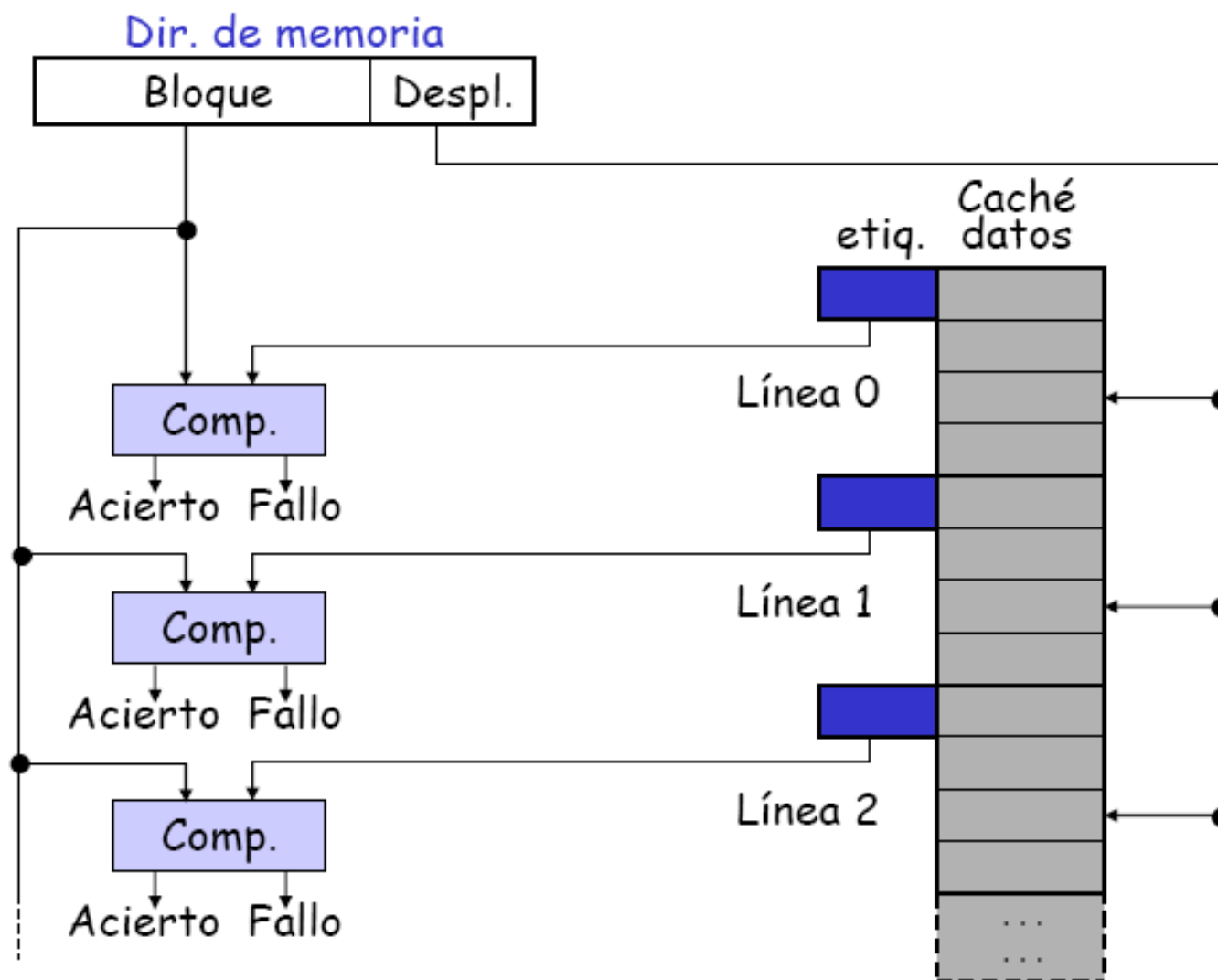


Política de
Sustitución

Dir. de memoria

Bloque

Despl.



Nuestro Ejemplo

Tamaño caché: 4 Kbytes
 Tamaño bloque: 4 bytes } → Caché de 1 Klínea de 4 bytes

Memoria principal: 64 Kbytes (16 Kbloques de 4 bytes)

Dirección M.P. Número de bloque (etiqueta) despl.

	14	2
--	----	---

Línea caché	Bloques de memoria principal
0	cualquiera
1	cualquiera
...	...
3FF	cualquiera

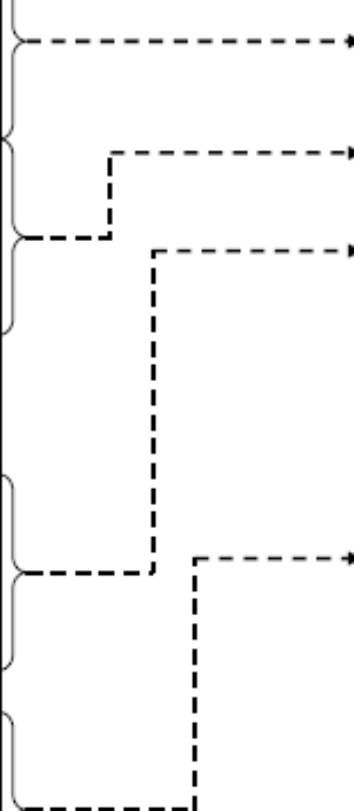
Direcciones de M.P.

Num. bloque		desp. datos
0000000000000000	00	13
0000000000000000	01	57
0000000000000000	10	92
0000000000000000	11	46
0000000000000001	00	11
0000000000000001	01	22
0000000000000001	10	33
0000000000000001	11	44
0000	...	...
...	...	...
0000111111111111	11	AB
0001000000000000	00	12
0001000000000000	01	23
0001000000000000	10	34
0001000000000000	11	45
...	...	...
1111111111111111	00	22
1111111111111111	01	22
1111111111111111	10	22
1111111111111111	11	22

Caché

Etiqu.	Datos	Línea
0000	13579246	000
0001	11223344	001
0400	12233445	002
...	...	...
334A	11223344	3FE
3FFF	22222222	3FF

14
32
10

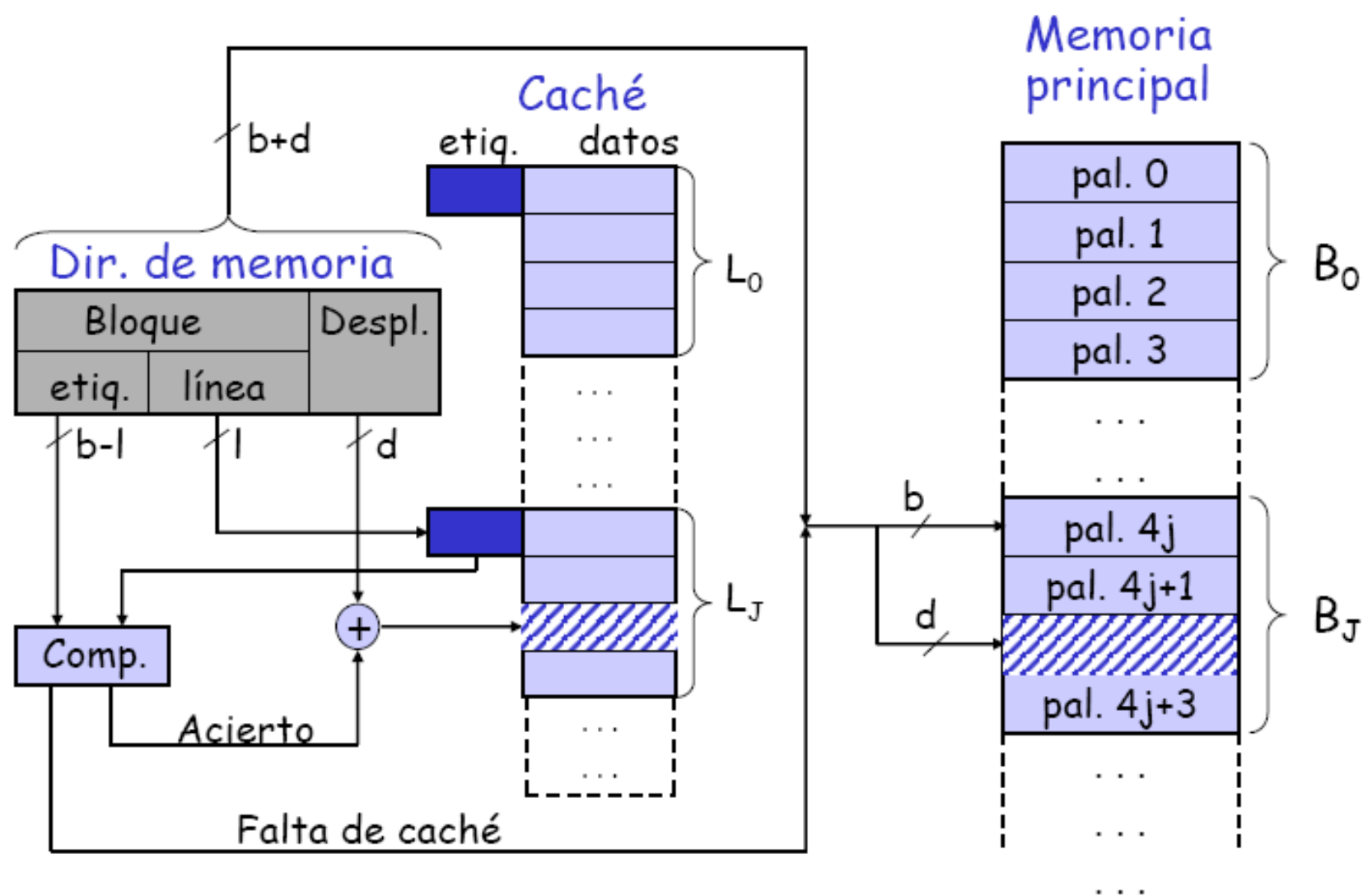


Dir. de memoria

Bloque		Despl.
etiq.	línea	

Caché de
L líneas

Línea caché	Bloques de memoria principal
0	0 , L , 2L , 3L , ...
1	1 , L+1 , 2L+1 , 3L+1 , ...
...	...
L-1	L-1 , 2L-1 , 3L-1 , 4L-1 , ...



Ejemplo

Tamaño caché: 4 Kbytes
 Tamaño bloque: 4 bytes

→ Caché de 1 Klínea de 4 bytes

Memoria principal: 64 Kbytes (16 Kbloques de 4 bytes)

Dirección M.P.	Número de bloque		despl.
	etiqueta	línea	
	4	10	2

Línea caché	Bloques de memoria principal
0	0, 400, 800, C00, ..., 3C00
1	1, 401, 801, C01, ..., 3C01
...	...
3FF	3FF, 7FF, BFF, ..., 3FFF

Direcciones de M.P.

Num. bloque desp. datos

0000	0000000000	00	13
0000	0000000000	01	57
0000	0000000000	10	92
0000	0000000000	11	46
0000	0000000001	00	11
0000	0000000001	01	22
0000	0000000001	10	33
0000	0000000001	11	44
0000	...	...	...
...	...	...	...
0000	1111111111	11	AB
0001	0000000000	00	12
0001	0000000000	01	23
0001	0000000000	10	34
0001	0000000000	11	45
...	...	...	...
1111	1111111111	00	22
1111	1111111111	01	22
1111	1111111111	10	22
1111	1111111111	11	22

...Ejemplo

Caché

Etiqu.	Datos	Línea
0000	13579246	000
0000	11223344	001
0110	87654321	002
...	...	...
1101	11223344	3FE
1111	22222222	3FF

← 4 ← 32 ← 10

¡ Hay Colisiones !

C_0	C_1
C_2	C_3

La caché se divide en C conjuntos de L líneas cada uno

Dir. de memoria

Bloque		Despl.
etiq.	conjunto	

El bloque se ubica en cualquier línea del conjunto

A muchos bloques les corresponderá el mismo conjunto

Política de sustitución dentro de cada conjunto

Nuestro Ejemplo

Tamaño caché: 4 Kbytes }
 Tamaño bloque: 4 bytes } → Caché de 1 Klínea de 4 bytes
Conjuntos de 4 bloques → Caché de 256 conjuntos
 Memoria principal: 64 Kbytes (16 Kbloques de 4 bytes)

Dirección
M.P.

Número de bloque		despl.
6	8	2
etiqueta		conjunto

Cjto. caché	Bloques de memoria principal
0	0, 100, 200, ..., 3F00
1	1, 101, 201, ..., 3F01
...	...
FF	FF, 1FF, 2FF, ..., 3FFF

Direcciones de M.P.

Num. bloque		desp. datos	
000000	00000000	00	13
000000	00000000	01	57
000000	00000000	10	92
000000	00000000	11	46
000000	00000001	00	11
000000	00000001	01	22
000000	00000001	10	33
000000	00000001	11	44
000000	...	...	...
...	...	...	...
000000	11111111	11	77
000001	00000000	00	AB
000001	00000000	01	12
000001	00000000	10	23
000001	00000000	11	34
...	...	...	...
111111	11111111	00	22
111111	11111111	01	22
111111	11111111	10	22
111111	11111111	11	22

CACHÉ

Etiqu.	Datos	Cjto.
000000	13579246	00
000001	AB122334	
010100	32454622	
010101	67524425	01
000000	11223344	
000110	ABCD1234	
000001	123ABC45	
100000	98712365	...
...	...	
...	...	
000001	11223344	FF
111111	22222222	
111000	99999999	
111001	88888888	

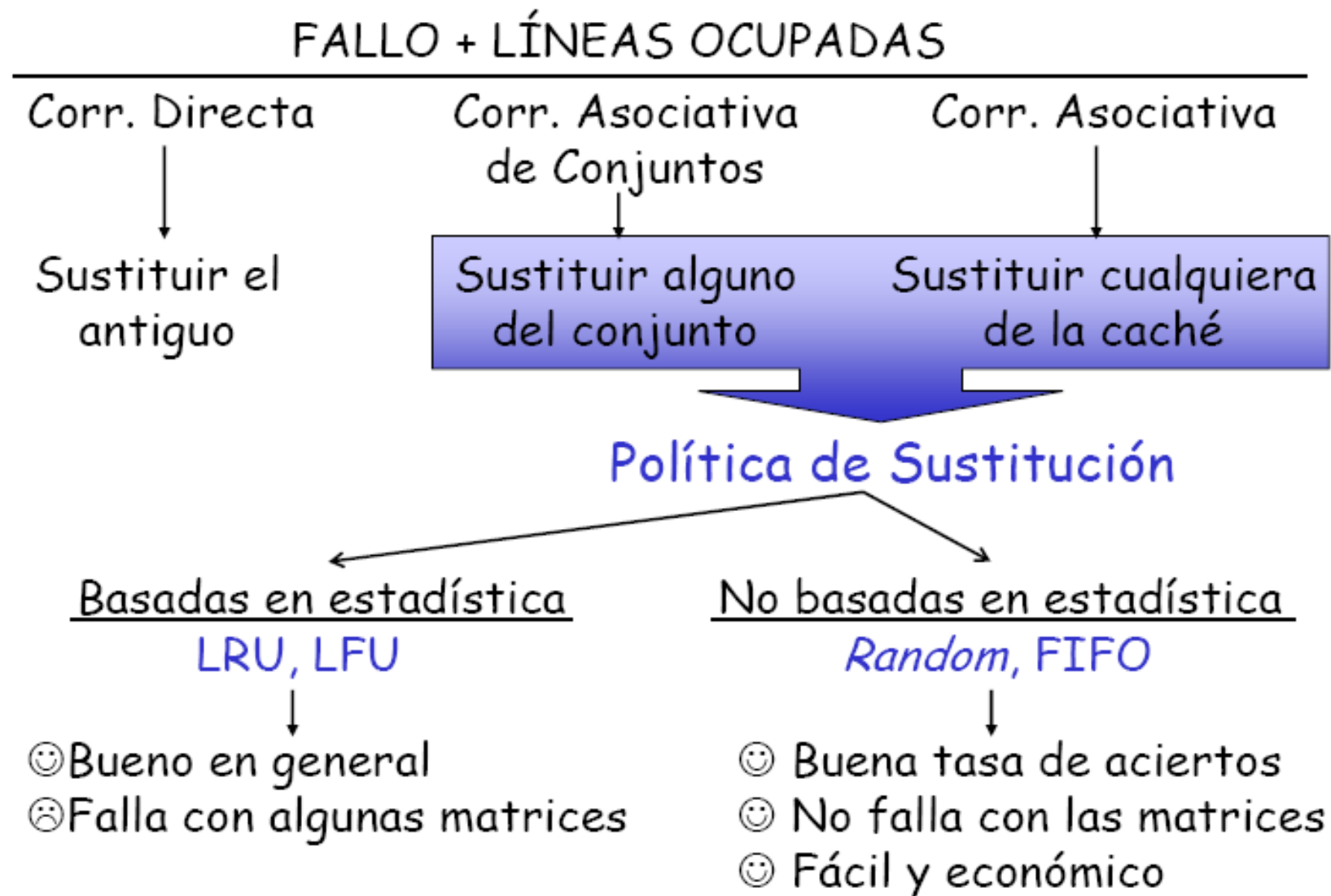
6

32

8

Comparación

- Directa
 - Un bloque de MP sólo puede estar en una línea
- Asociativa
 - Un bloque puede estar en cualquier línea
- Asociativa por conjuntos
 - Un bloque puede almacenarse en un conjunto de líneas



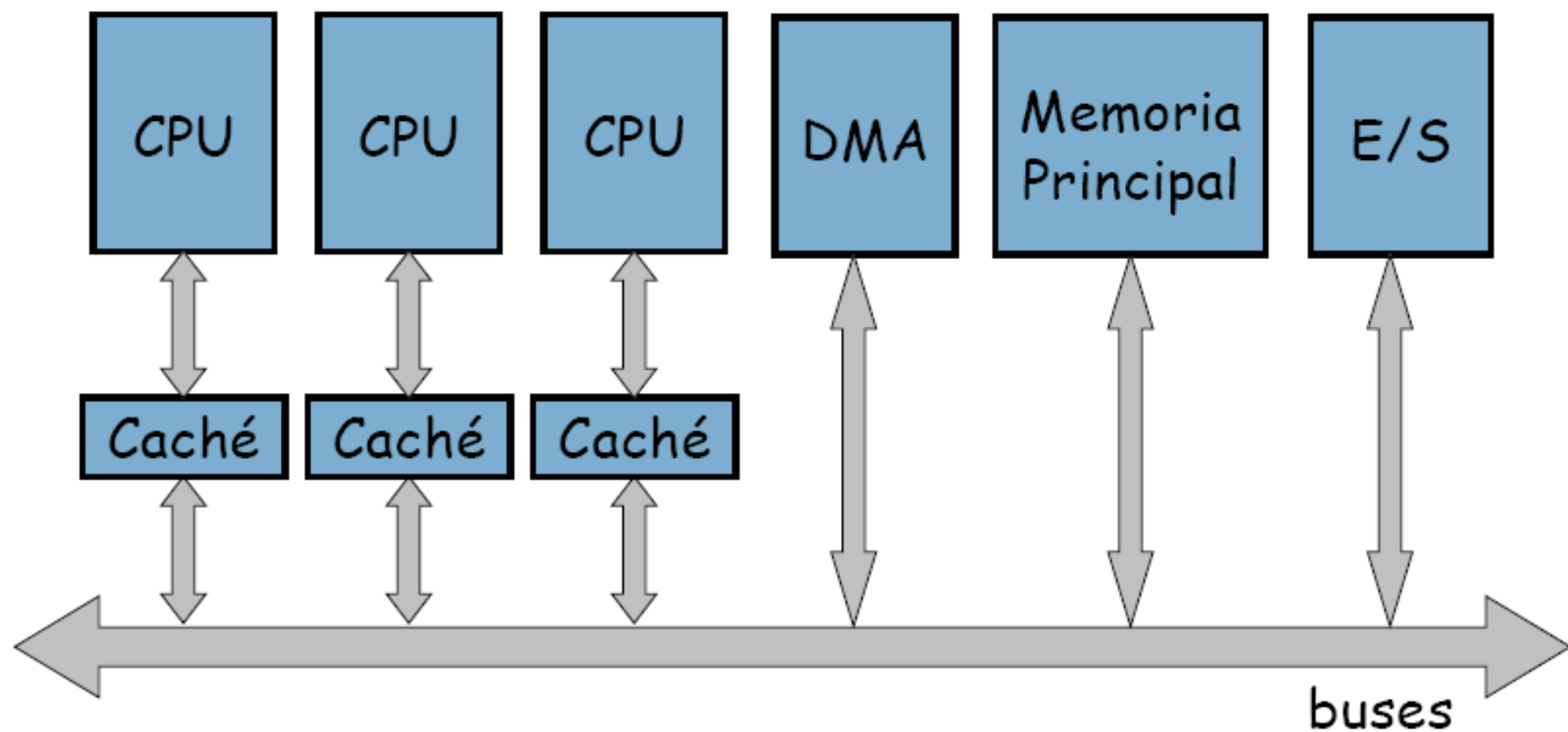
Memoria Asociativa de Conjuntos						
Tamaño	Dos Vías		Cuatro Vías		Ocho Vías	
	LRU	Random	LRU	Random	LRU	Random
16 Kb	5,18%	5,69%	4,67%	5,29%	4,39%	4,96%
64 Kb	1,88%	2,01%	1,54%	1,66%	1,39%	1,53%
256 Kb	1,15%	1,17%	1,13%	1,13%	1,12%	1,12%

*Comparación de tasa de fallos
según diversos tamaños y políticas de sustitución*

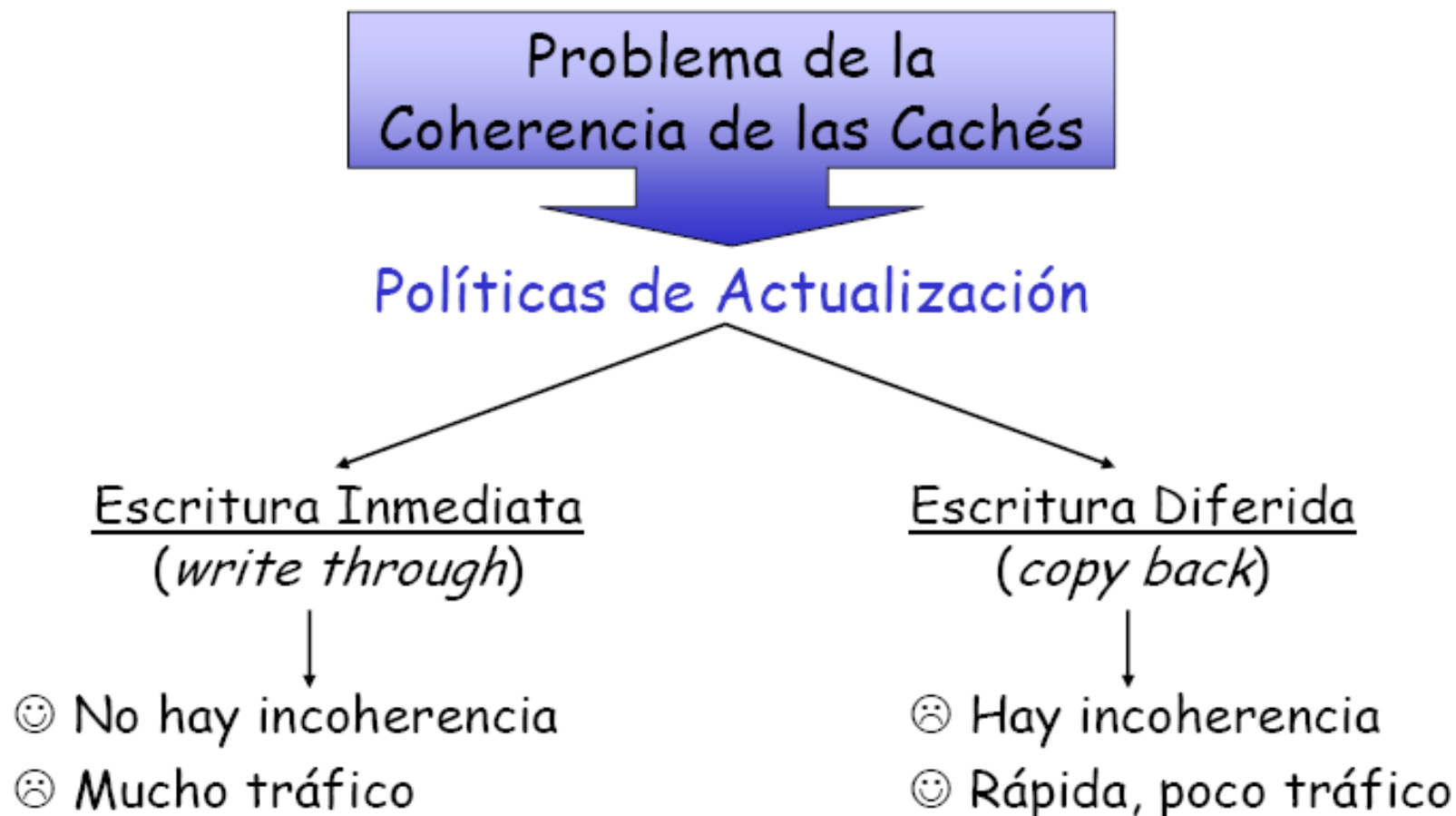
Cuestiones ante
una Operación
de Escritura

❏ Problema de la Coherencia
de las Cachés

❏ Falta de Caché en Escritura

El Problema de Coherencia de las Cachés

Ante una escritura...



Ante una escritura...

¿ Falta de Caché en Escritura ?

Dos Opciones

Escritura sin asignación

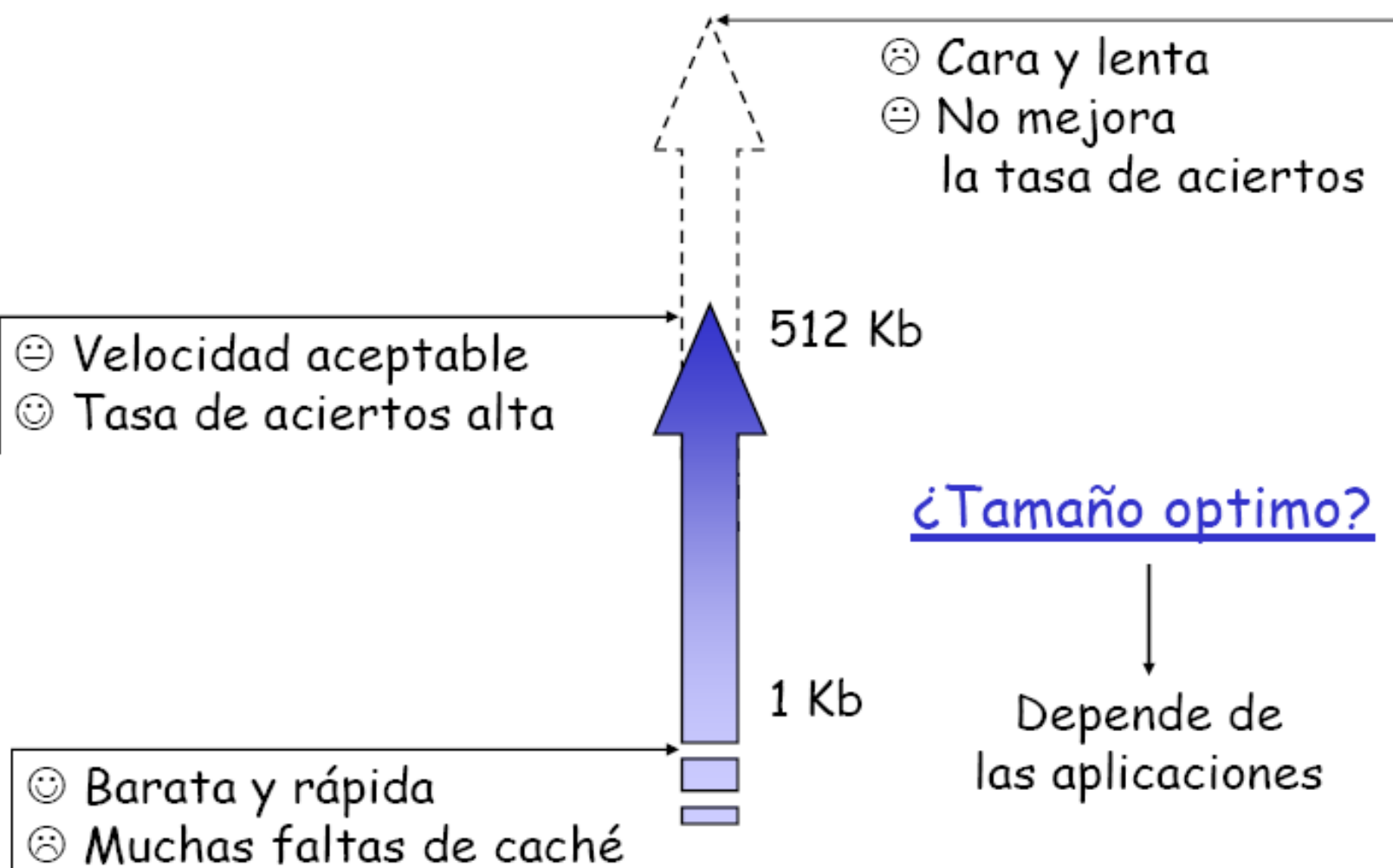
Escribir directamente
en la memoria

Escritura con asignación

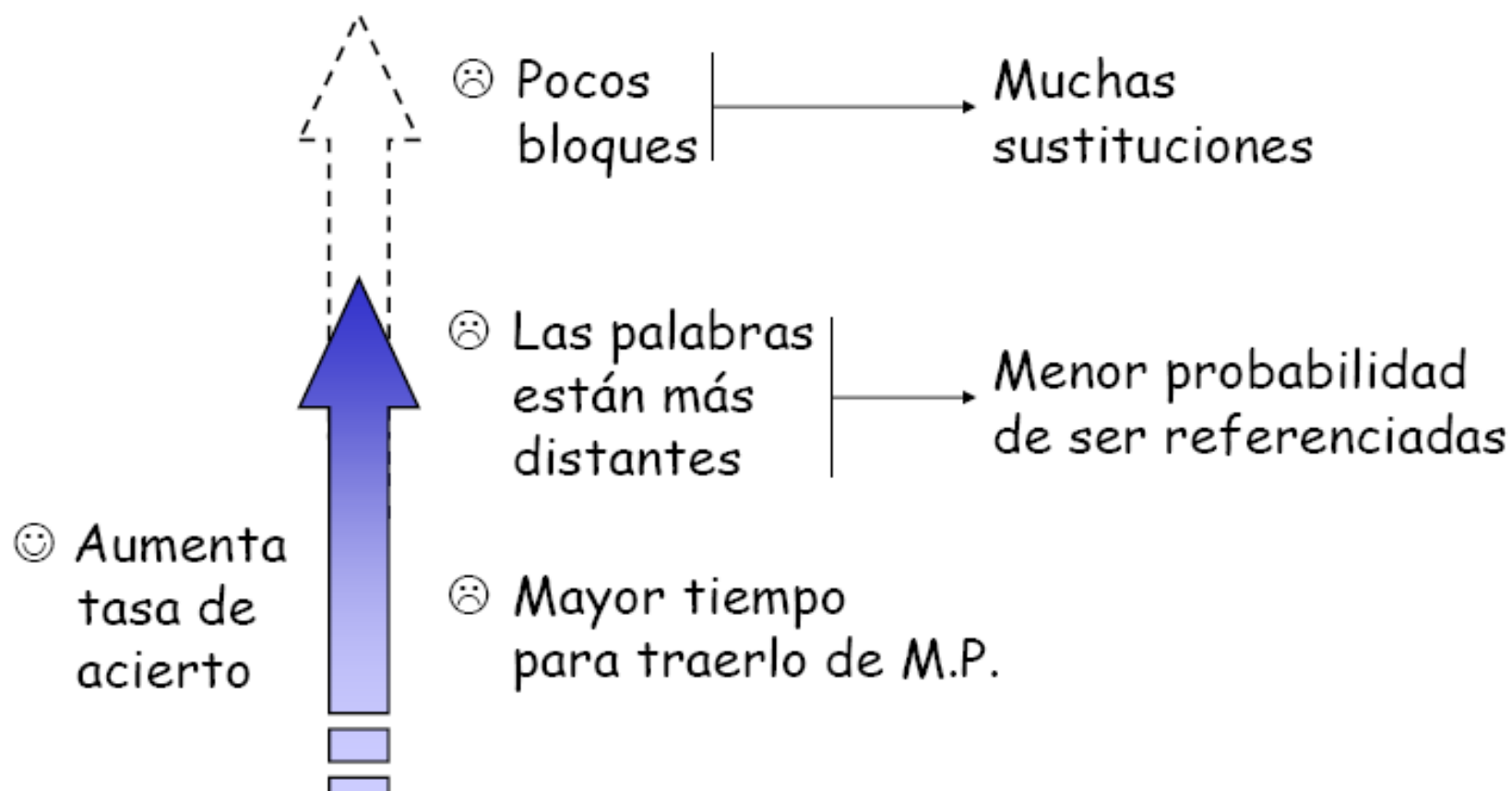
1º Traer el dato a caché
2º Escribir en la caché

Escritura diferida → Escritura con asignación

Escritura inmediata → Escritura sin asignación



¡ Probabilidad de acceso al mismo bloque < Probabilidad de acceso a un bloque reemplazado !

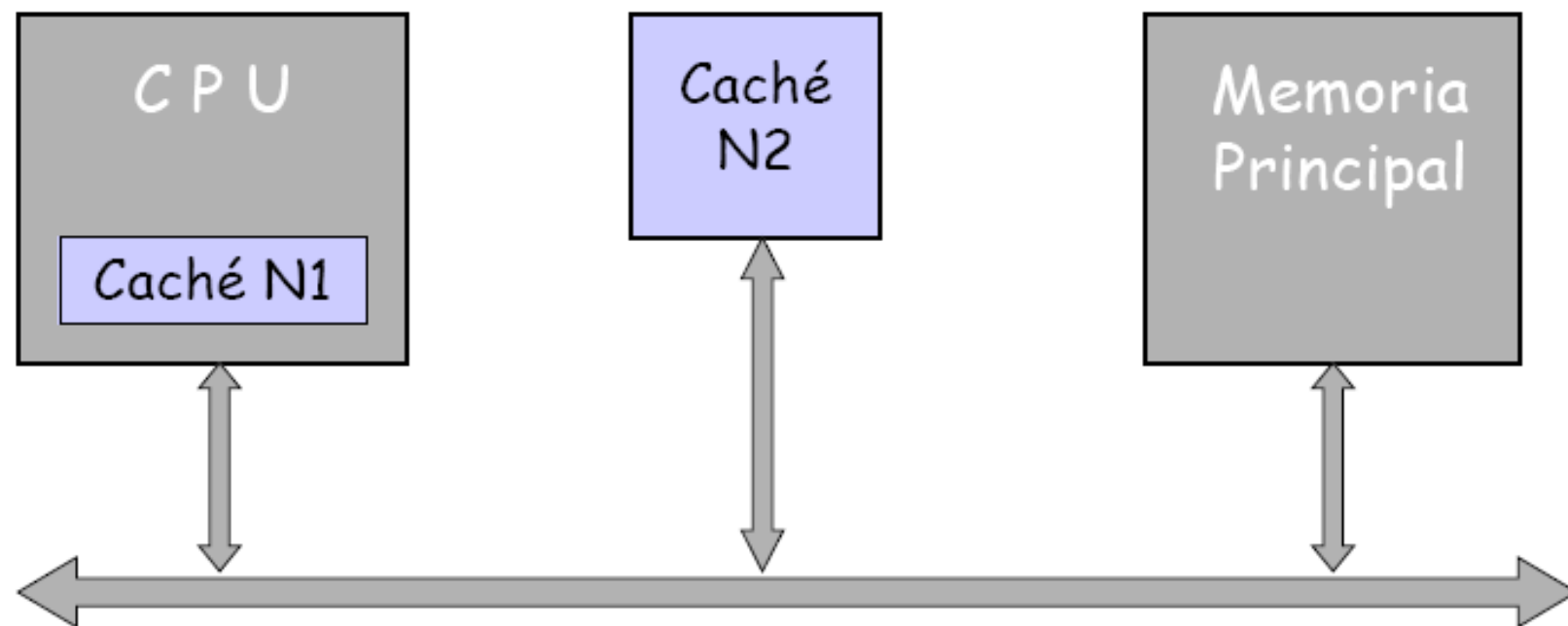


Tamaño bloque	Tamaño de la Caché				
	1K	4K	16K	64K	256K
16	15,05%	8,57%	3,94%	2,04%	1,09%
32	<u>13,34%</u>	7,24%	2,87%	1,35%	0,70%
64	13,76%	<u>7,00%</u>	<u>2,64%</u>	1,06%	0,51%
128	16,64%	7,78%	2,77%	<u>1,02%</u>	<u>0,49%</u>
256	22,01%	9,51%	3,29%	1,15%	0,49%

*Tasa de faltas de caché, según el tamaño de bloque,
para distintos tamaños de caché*

Tamaño bloque	Tiempo por falta	Tamaño de la Caché				
		1K	4K	16K	64K	256K
16	42	7,321	4,599	2,655	1,859	1,458
32	44	<u>6,870</u>	<u>4,186</u>	<u>2,263</u>	1,594	1,308
64	48	7,605	4,360	2,267	<u>1,509</u>	<u>1,245</u>
128	56	10,318	5,357	2,551	1,571	1,274
256	72	16,847	7,847	3,369	1,828	1,353

Tiempo medio de acceso a memoria, según el tamaño de bloque, para distintos tamaños de caché



Gran ventaja de la caché N1
DEJA LIBRE EL BUS

En los programas hay
dos grandes áreas

- Instrucciones
- Datos

¿Una caché compartida o dos separadas?

COMPARTIDA

- ☺ Equilibrio automático
de instrucciones y datos

↓
Mayor tasa de aciertos

- ☺ Más simple de diseñar
e implementar

SEPARADAS

(Arquitectura Harvard)

- ☺ En procesadores con
prefetching y *pipeline* se
permite acceso a datos
mientras se extrae la
siguiente instrucción

Ejemplo Caché Actual

The screenshot shows the CPU-Z application window. The 'CPU' tab is selected. The 'Processor' section displays the following information:

Name	Intel Core i7 2670QM		
Code Name	Sandy Bridge	Max TDP	45.0 W
Package	Socket 988B rPGA		
Technology	32 nm	Core VID	1.126 V
Specification	Intel(R) Core(TM) i7-2670QM CPU @ 2.20GHz		
Family	6	Model	A
Ext. Family	6	Ext. Model	2A
Stepping	7	Revision	D2
Instructions	MMX, SSE, SSE2, SSE3, SSSE3, SSE4.1, SSE4.2, EM64T, VT-x, AES, AVX		

The 'Cache' section displays the following information:

L1 Data	4 x 32 KBytes	8-way
L1 Inst.	4 x 32 KBytes	8-way
Level 2	4 x 256 KBytes	8-way
Level 3	6 MBytes	12-way

The 'Clocks (Core #0)' section displays the following information:

Core Speed	2194.76 MHz
Multiplier	x 22.0 (8 - 31)
Bus Speed	99.76 MHz
Rated FSB	

The 'Selection' dropdown is set to 'Processor #1'. The 'Cores' field shows 4 and the 'Threads' field shows 8.

At the bottom, the CPU-Z logo and version 'Ver. 1.73.0.x64' are displayed, along with 'Tools', 'Validate', and 'Close' buttons.

Ejemplo Caché Actual

intel.com/es-es/products/53469/Intel-Core-i7-2670QM-Processor-6M-Cache-up-to-3_10-GHz

CS 571, Operating Syst... Acceder NMEA data Configurator de NI De... ARQII_08-MIMD_A4x6... O06G150V01401 - TEMA

Aprender y desarrollar Soporte

intel

Menú ARK ▶

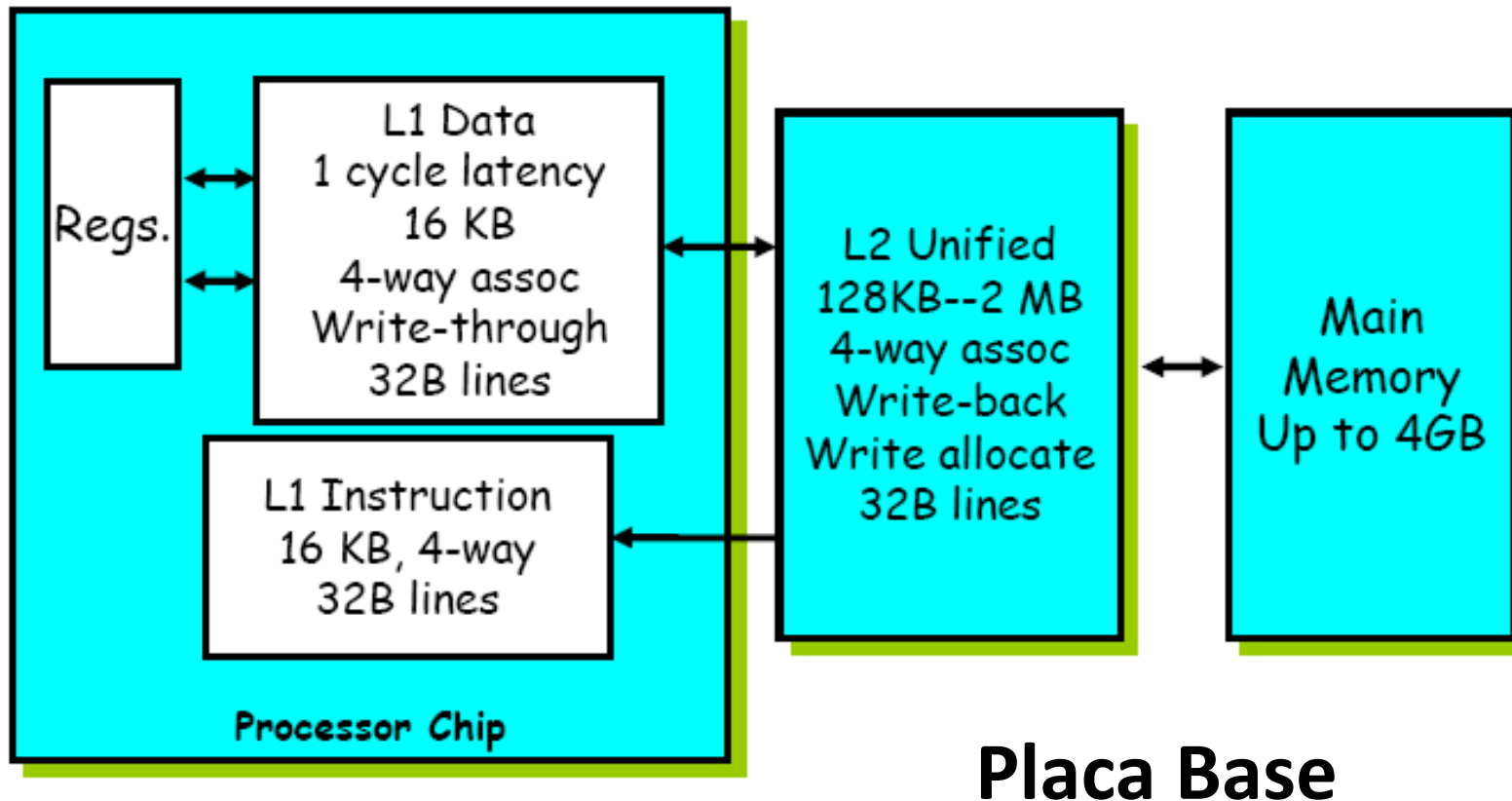
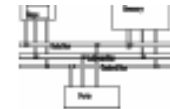
Escriba aquí

Especificaciones de memoria	Fecha de lanzamiento	Q4'11
Especificaciones gráficas	Litografía	32 nm
Opciones de expansión	Precio de cliente recomendado	N/A
Especificaciones del paquete		
Tecnologías avanzadas		
Tecnología Intel® para protección de datos		
Tecnología Intel® de protección de plataforma		
Productos compatibles		
Imágenes del producto		

- Rendimiento	
Núcleos	4
Nº de subprocesos	8
Frecuencia base del procesador	2.20 GHz
Frecuencia turbo máxima	3.10 GHz
Caché	6 MB SmartCache
Velocidad del bus	5 GT/s DMI
TDP	45 W

Ejemplo Caché No Actual

Intel Pentium III cache hierarchy



Bibliografía

- <http://www.houstonisd.org/cms/lib2/tx01001591/centricity/domain/10450/chap6.pdf>
- <http://www.hardwaresecrets.com/category/memory/>
- <http://www.jedec.org/>
- http://www.dia.eui.upm.es/Asignatu/arq_com/Paco/7-Cache.pdf

