



UNIVERSIDAD
UPP

Tamaño del caché-Funciones de mapeo-Mapeo

Materia:

Diseño de Lenguajes

Alumno:

Elí Noé Herrera Loría

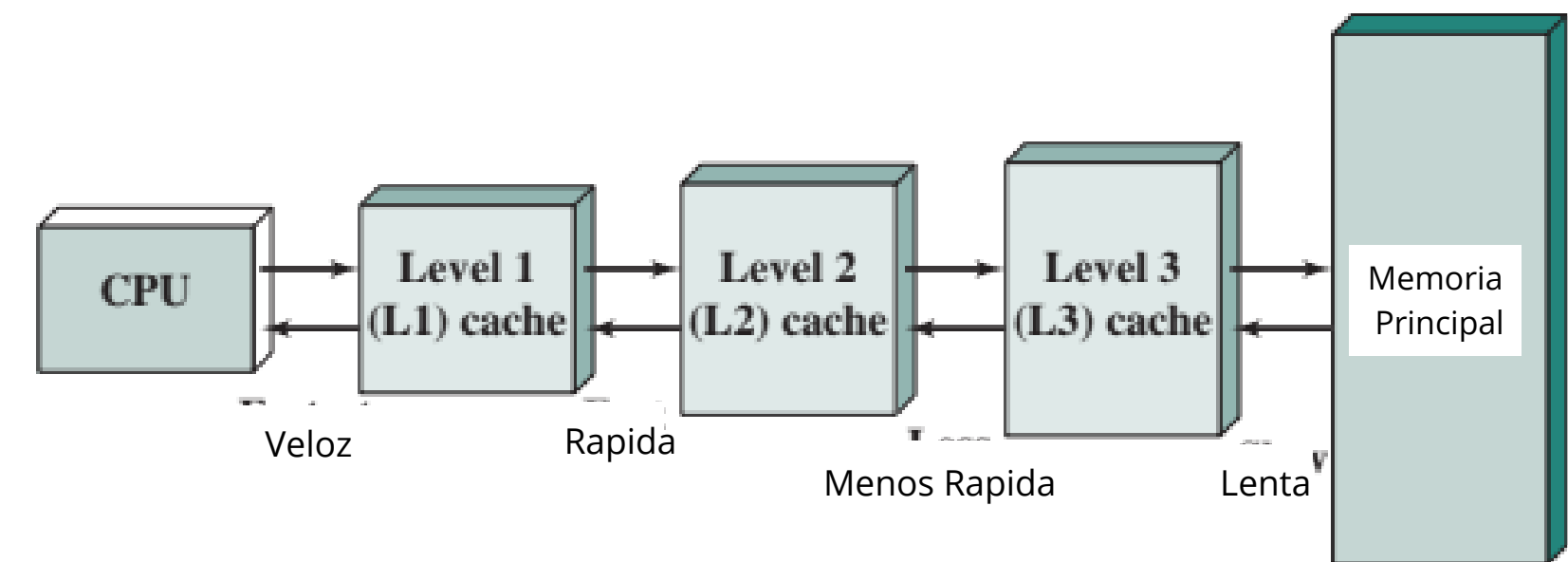
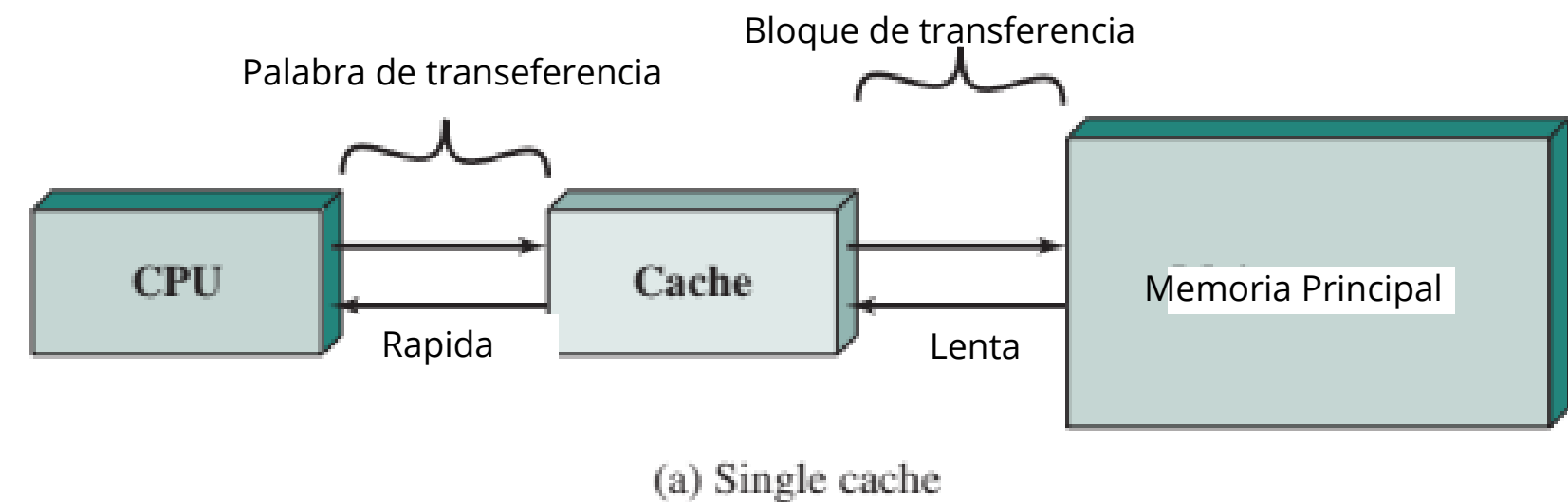
Docente:

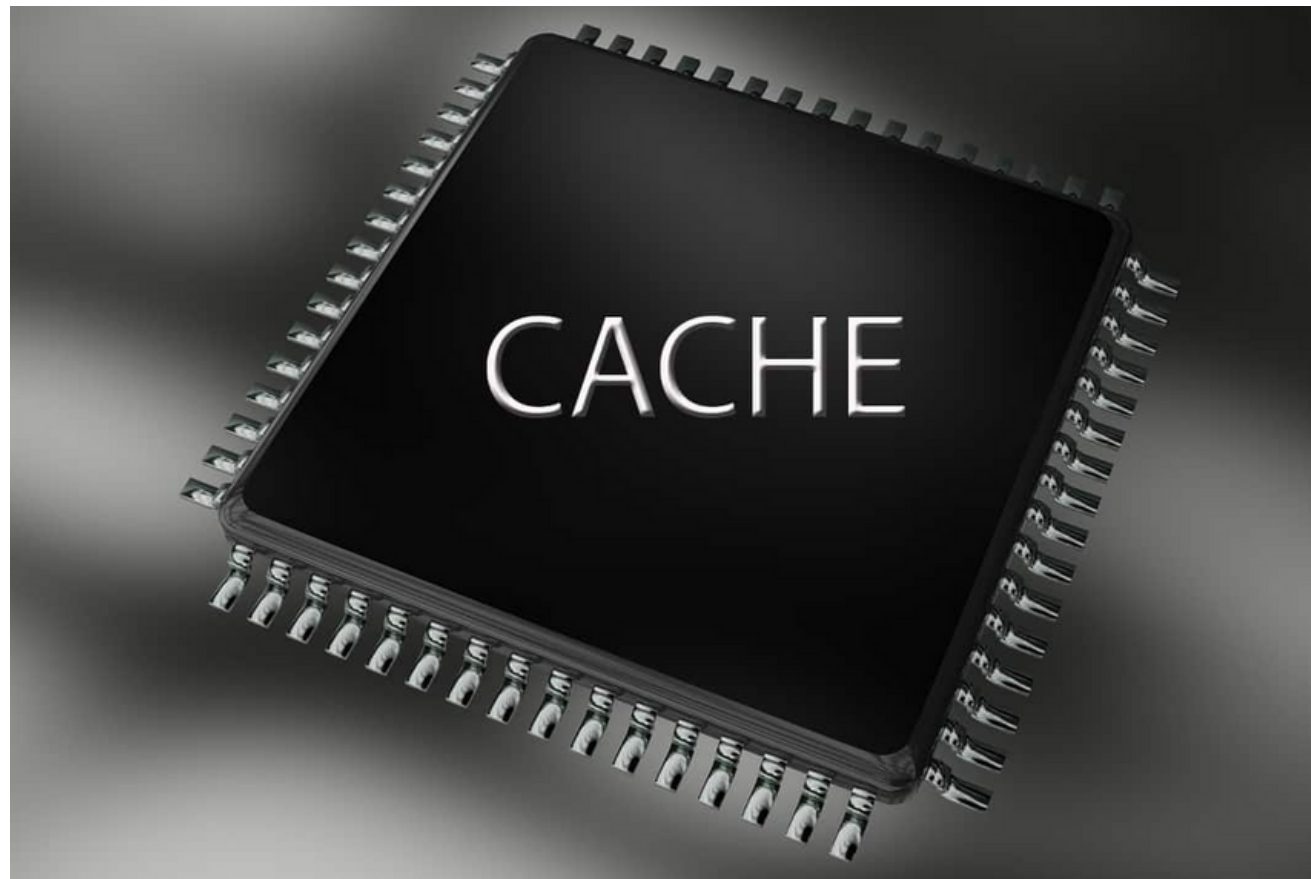
M. en C. Jorge J. Pedrozo Romero

Licenciatura en Desarrollo en Software

¿Qué es la memoria caché?

En informática, una memoria caché es una capa de almacenamiento de datos de alta velocidad que almacena un subconjunto de datos, normalmente transitorios, de modo que las solicitudes futuras de dichos datos se atienden con mayor rapidez que si se debe acceder a los datos desde la ubicación de almacenamiento principal.





Tamaño de la caché

El tamaño de la caché es un factor importante, pues este determinará la rapidez de nuestros mapeos.

Cuanto más pequeña la memoria, más rápida será, pues los datos estarán reducidos y por ende, el tiempo de mapeo es menor. Claro que el tamaño de la caché depende en gran medida para que será usada.

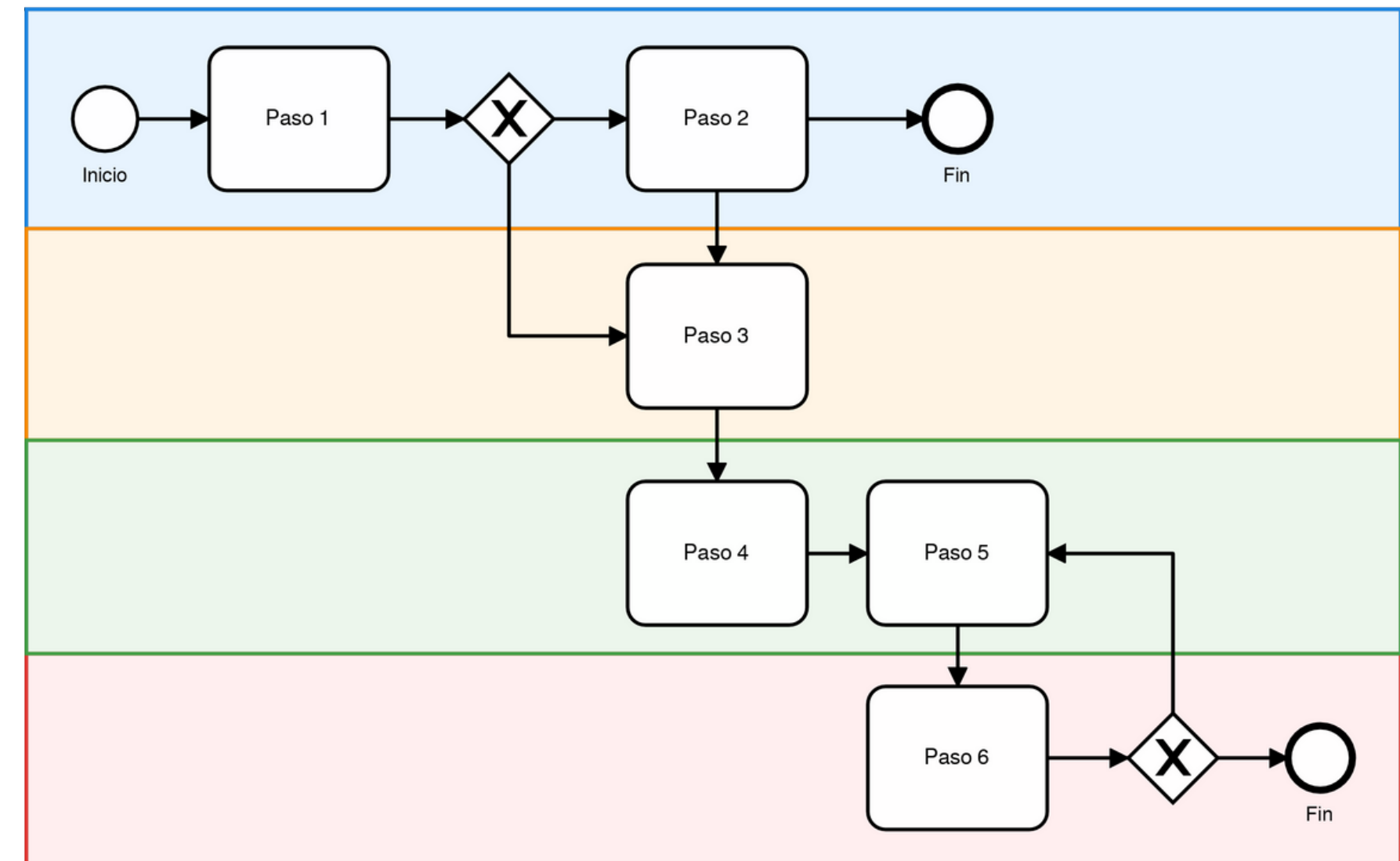
Tamaño ideal

Ahora, el tamaño ideal de la caché no existe.

¿Por qué? Porque al ser la caché muy sensible a la carga de trabajo no se puede llegar a ese tamaño ideal, aunque este puesto en el mismo lugar del chip

¿Qué es el mapeo?

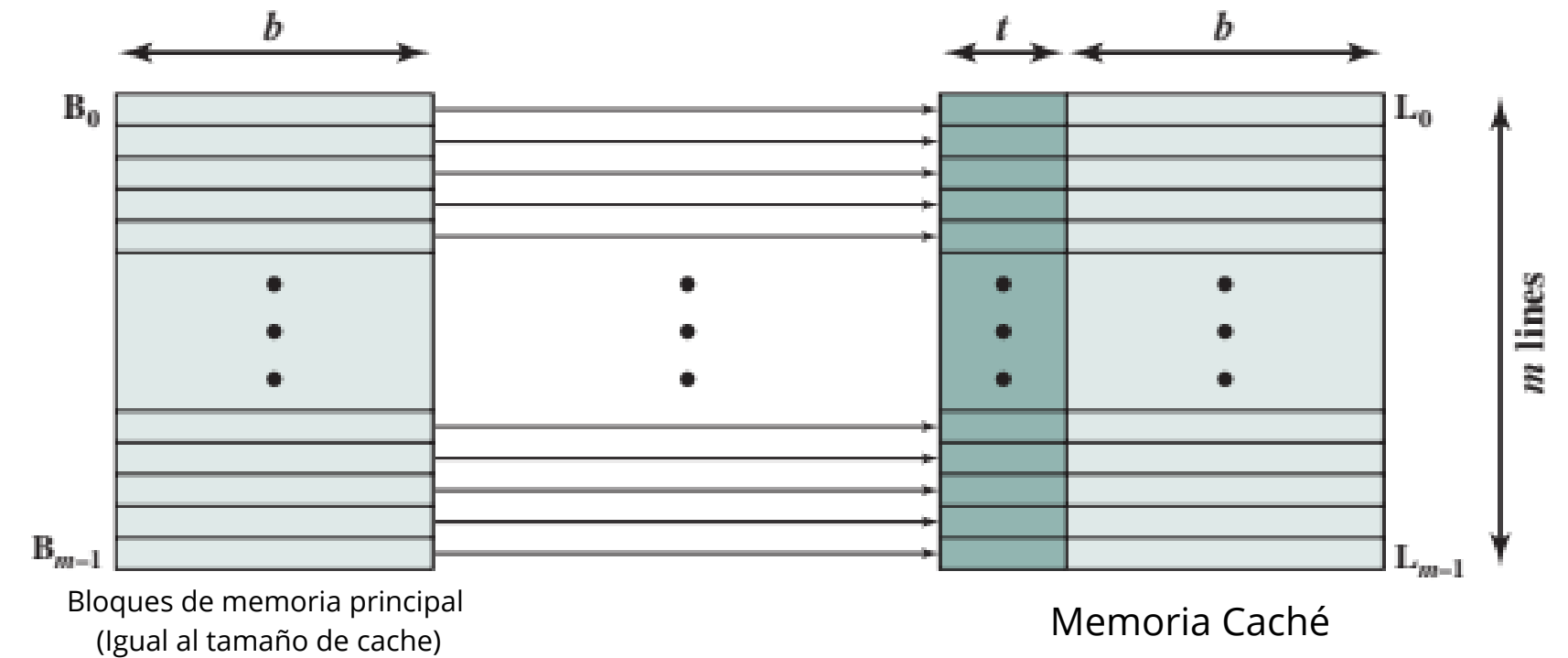
El mapeo se refiere al proceso de crear representaciones visuales de información o datos para proporcionar una comprensión más clara y estructurada



Funciones de Mapeo

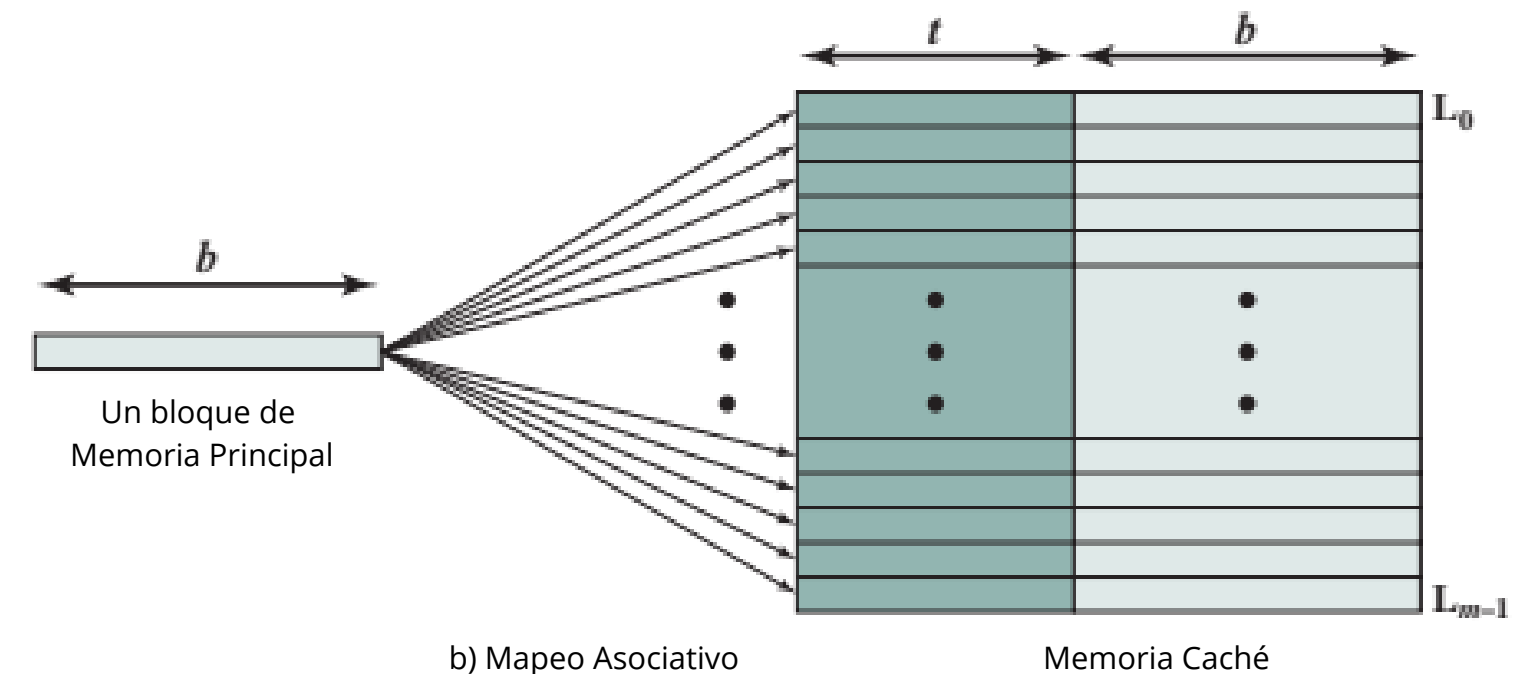
Dado que las líneas de caché son menos de los bloques de Memoria Principal se debe tener algún algoritmo para asignar la línea a un bloque de la Memoria Principal.

Se pueden utilizar tres técnicas: directa, asociativa y asociativa por conjuntos.



a) Mapeo Directo

b = largo del bloque en bits
 t = largo de la etiqueta en bits



b) Mapeo Asociativo

Mapeo Directo

El mapeo directo es un método de almacenar información para un fácil acceso en una computadora. En este enfoque, la información utilizada recientemente se almacena en un caché para que la computadora pueda encontrar rápidamente la información la próxima vez que la necesite¹².

Aquí está cómo funciona:

Asignación de Líneas en la Caché:

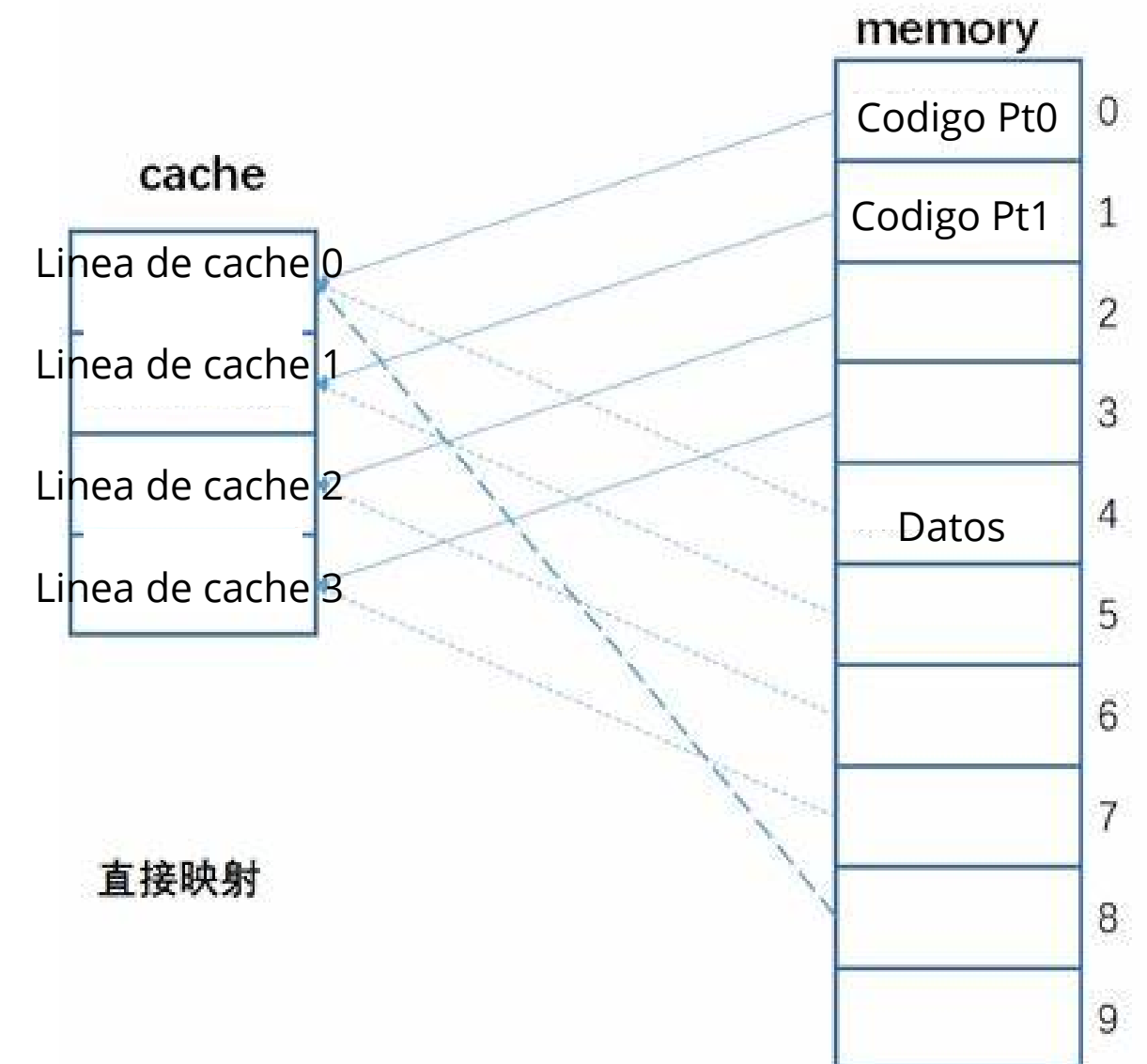
Cada bloque de datos en la memoria principal se asigna a una línea específica en la caché.

Dado que la caché es más pequeña que la memoria principal, varios bloques comparten una sola línea en la caché.

Sobreescritura Constante:

Los datos en la caché se sobrescriben constantemente a medida que se necesitan nuevos datos.

Si una línea ya está llena cuando se necesita escribir un nuevo bloque, se sobrescribirá un bloque antiguo.



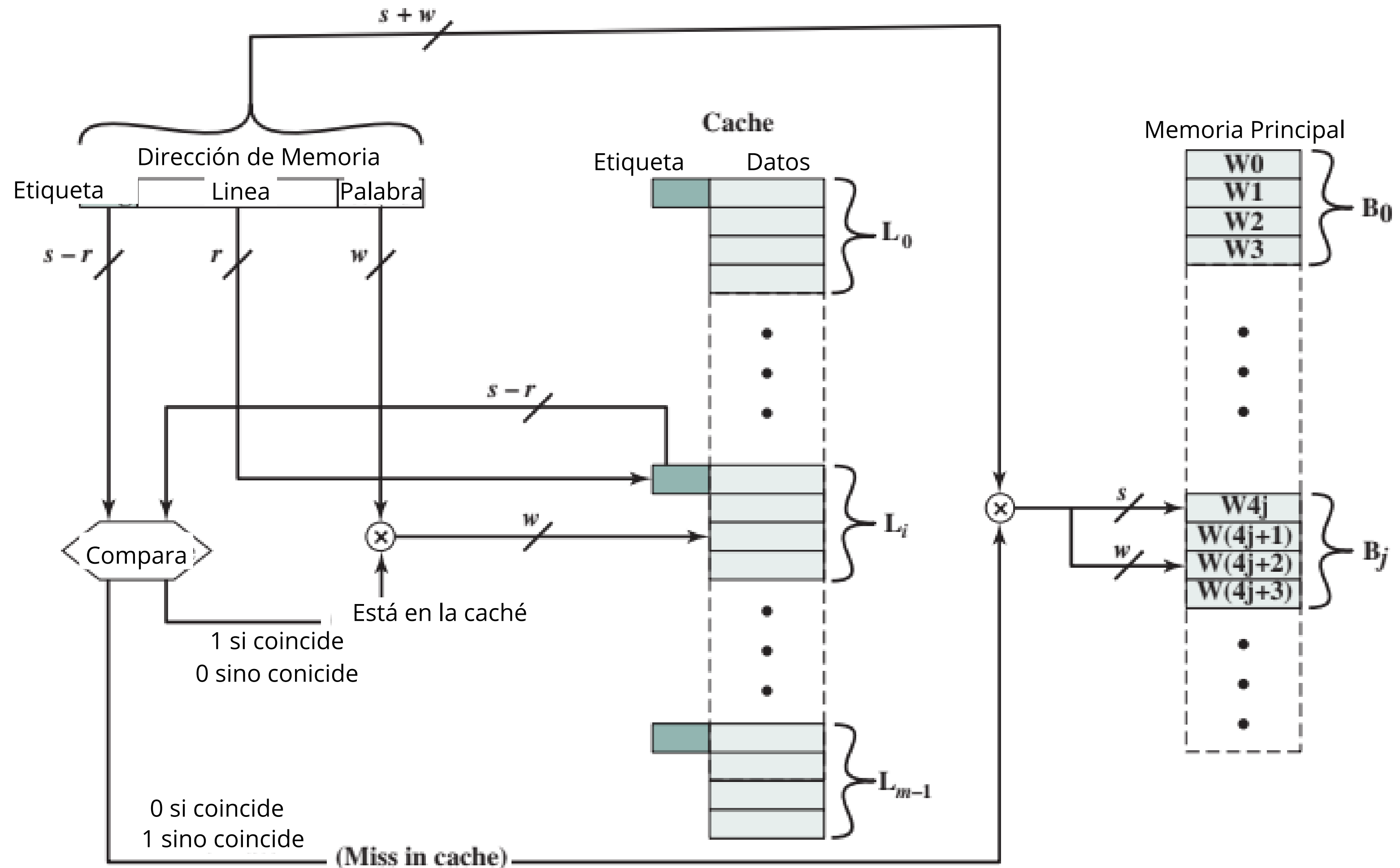


Figure 4.9 Direct-Mapping Cache Organization

Ejemplo 4.2

Para los tres casos, el ejemplo incluye los siguientes elementos:

- La caché puede contener 64 kB.
- Los datos se transfieren entre la memoria principal y la caché en bloques de 4 bytes cada uno. Esto significa que la caché está organizada como $16K = 2^{14}$ líneas de 4 bytes cada una.
- La memoria principal consta de 16 MB, y cada byte se puede direccionar directamente mediante una dirección de 24 bits ($2^{24} = 16\text{ M}$). Por tanto, a efectos de mapeo, podemos considerar que la memoria principal consta de bloques de 4M de 4 bytes cada uno.

Ejemplo

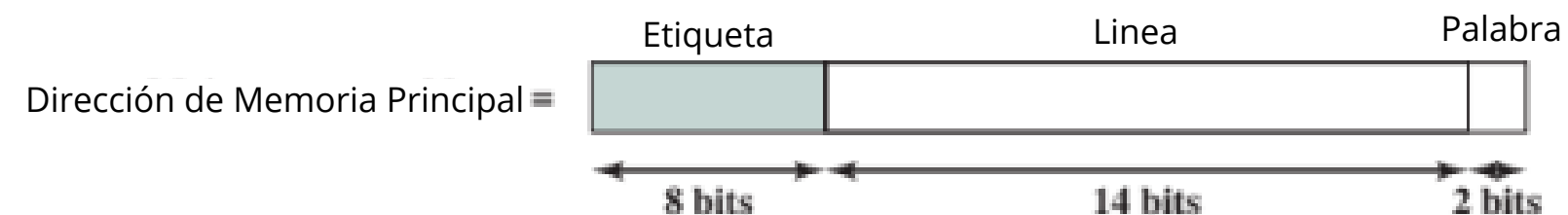
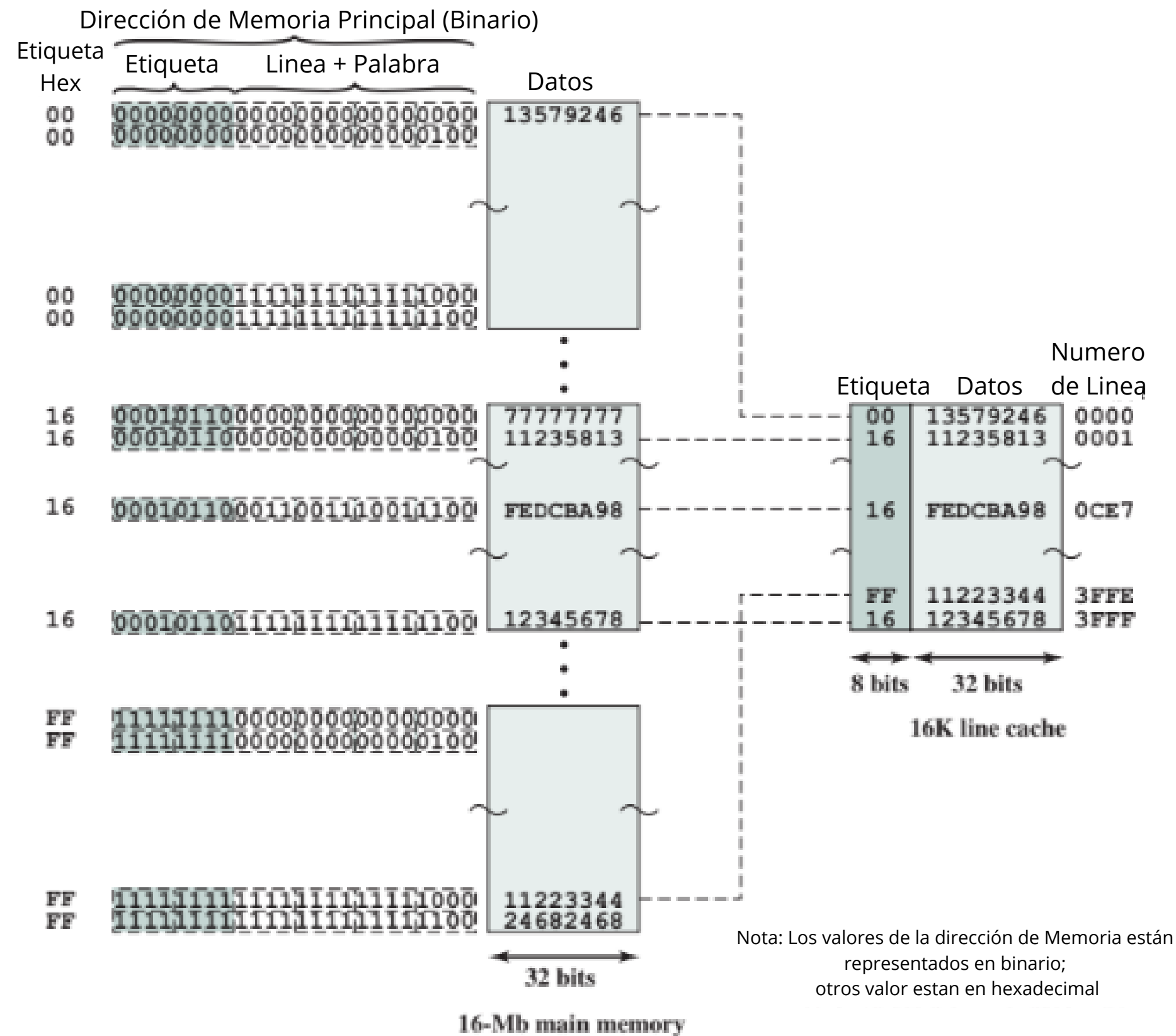


Figure 4.10 Direct Mapping Example

Ejemplo 4.2a

Muestra nuestro sistema de ejemplo usando mapeo directo.

$$m = 16K = 2^{14}$$

$$i = j \text{ modulo } 2^{14}$$

. El mapeo se convierte en:

Cache Line	Starting Memory Address of Block
0	000000, 010000, ..., FF0000
1	000004, 010004, ..., FF0004
⋮	⋮
$2^{14} - 1$	00FFFC, 01FFFC, ..., FFFFFC

Tenga en cuenta que no hay dos bloques que se asignen al mismo número de línea que tengan el mismo número de etiqueta. Por lo tanto, los bloques con direcciones iniciales 000000, 010000, ..., FF0000 tienen números de etiqueta 00, 01, ..., FF, respectivamente.

Volviendo a la Figura 4.5, una operación de lectura funciona de la siguiente manera. El sistema de caché se presenta con una dirección de 24 bits. El número de línea de 14 bits se utiliza como índice en el caché para acceder a una determinada línea.

Si el número de etiqueta de 8 bits coincide con el número de etiqueta actualmente almacenado en esa línea, entonces el número de palabra de 2 bits se usa para seleccionar uno de los 4 bytes en esa línea. De lo contrario, el campo etiqueta más línea de 22 bits se utiliza para recuperar un bloque de la memoria principal. La dirección real que se utiliza para la búsqueda es la etiqueta más línea de 22 bits concatenada con dos bits 0, de modo que se recuperan 4 bytes comenzando en un límite de bloque.