

Memoria Cache

Memoria Caché – Ejemplo comercial

- Procesador Intel® Core™ i7 – 3960X
- 6 Núcleos
- Velocidad de clock 3,3 GHz
- 15 MB de SmartCache

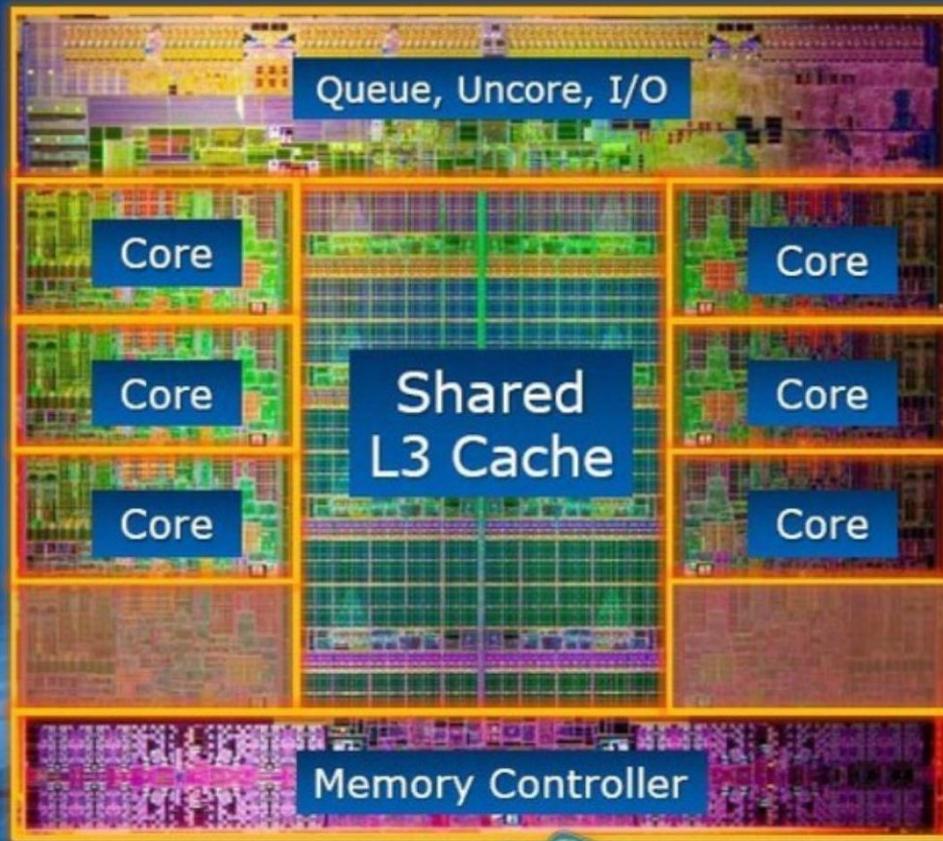
Caché



El Caché de CPU es un área de memoria rápida ubicada en el procesador. El Caché inteligente Intel® se refiere a la arquitectura que permite a todos los núcleos compartir dinámicamente el acceso al caché de alto nivel.

Memoria Caché – Ejemplo comercial

Intel® Core™ i7-3960X Processor Die Detail



Memoria Caché - Causa

- Los programas se ejecutan en pasos secuenciales
- Las variables se alojan en zonas adyacentes
- Recordar un bucle en ASM:

```
ciclo_a:
```

```
    mov    EAX,EBP  
    mov    [EBP+EAX-24],88  
    inc    [EBP-8]
```

```
ciclo_b:
```

```
    cmp    [EBP-8],15  
    jle    ciclo_a
```

Memoria Caché - Tecnología

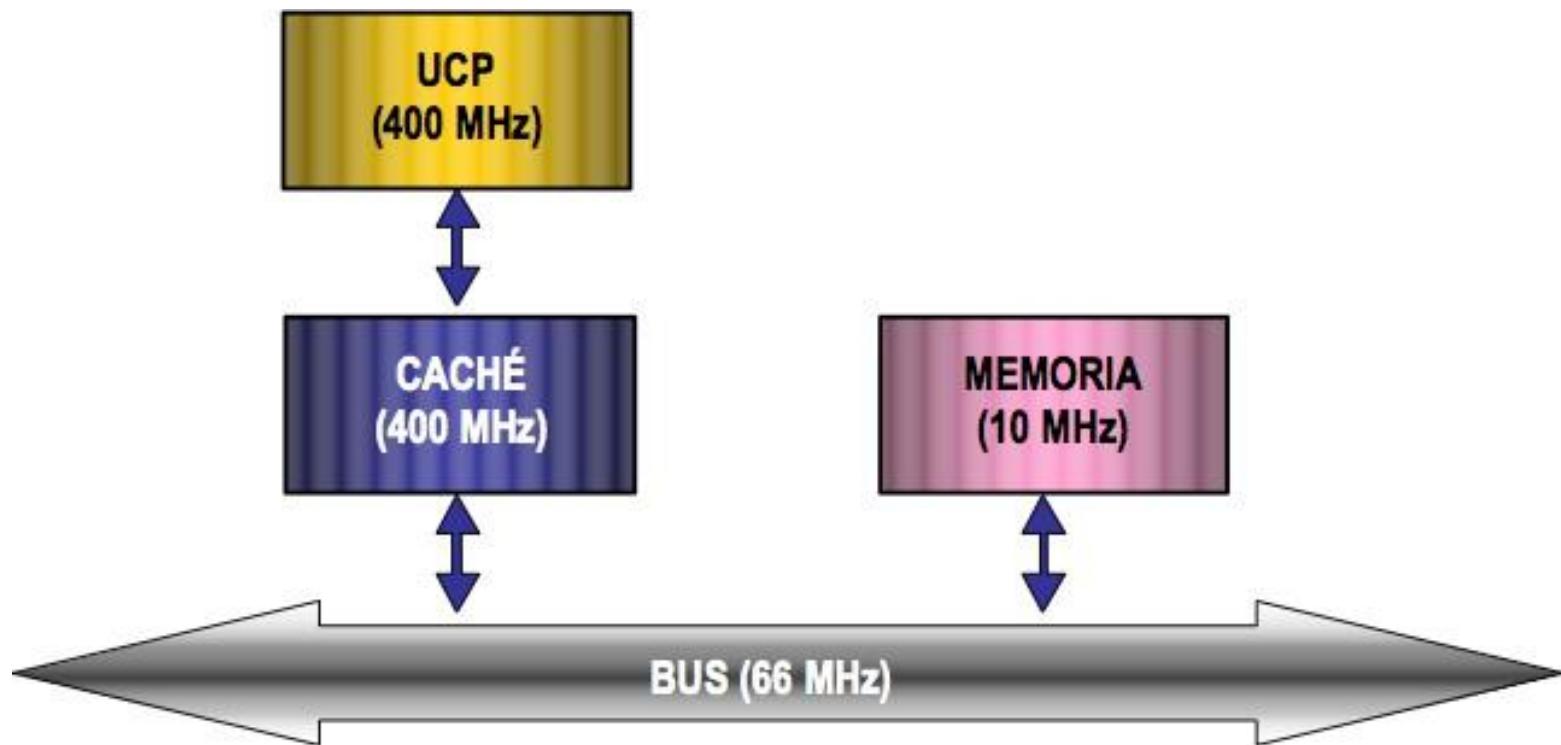
- Memoria DRAM mas económica y lenta
- Tecnología SRAM (Static RAM)
- Velocidad de respuesta acorde al CPU
- Guarda copias de la DRAM

Memoria Caché

- Diferencias de velocidades
 - Cuello de botella
 - Pérdida de performance

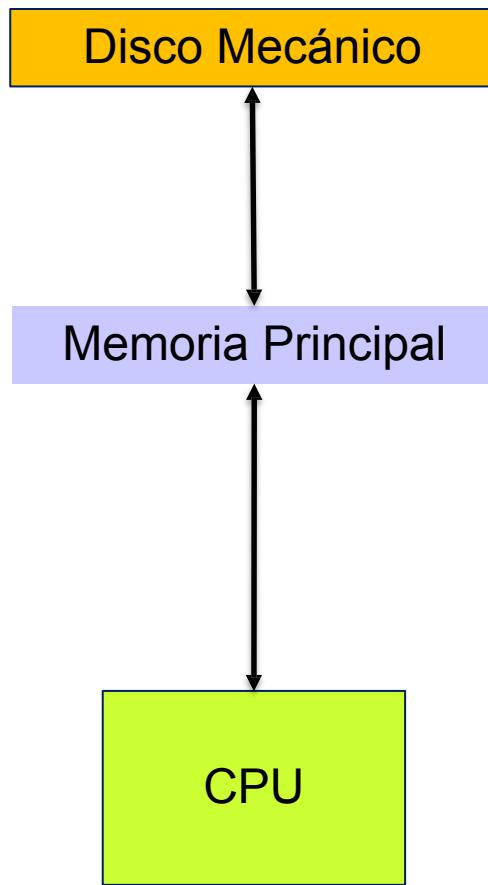


Memoria Caché - Funcionamiento

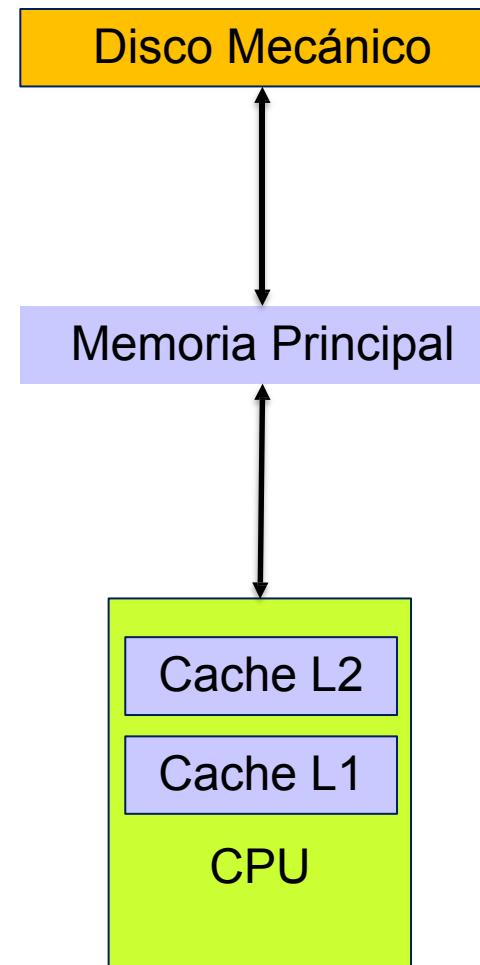


Memoria Caché - Evolución

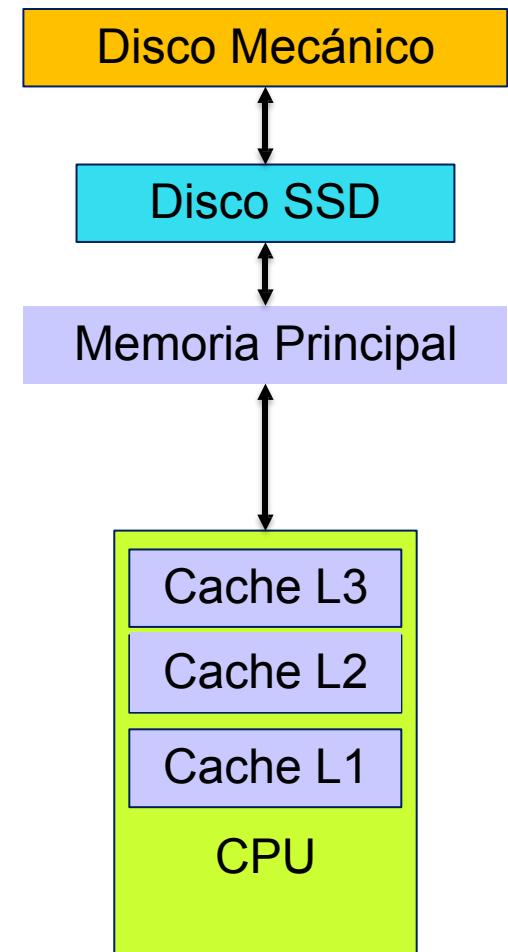
Hasta los 80's



90's y 2000's



2010



Comparativa disco-ram

Memory lookup vs SSD read

slowdown from real time



28123275x slowdown

cpu



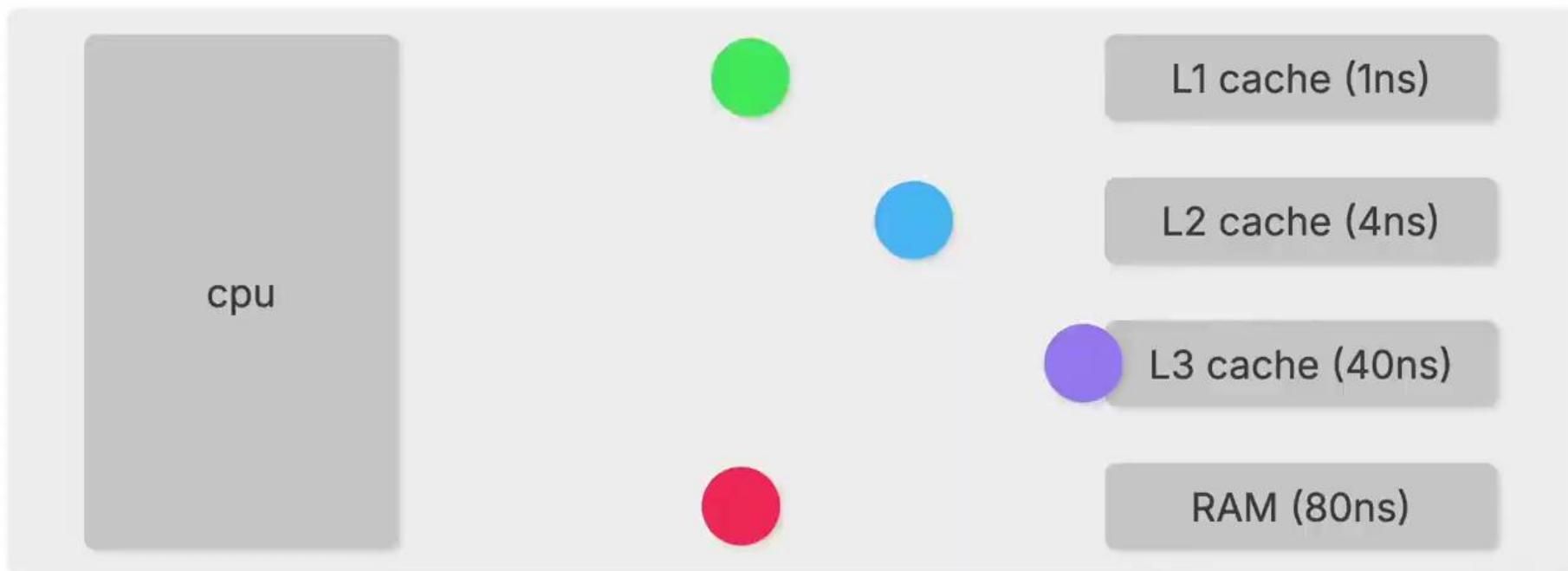
memory (50ns)



SSD (100μs)

Comparativa niveles de cache

CPU Cache and RAM



Comparativa “humana”

Event	Latency	Scaled
1 CPU cycle	0.3 ns	1 s
Level 1 cache access	0.9 ns	3 s
Level 2 cache access	2.8 ns	9 s
Level 3 cache access	12.9 ns	43 s
Main memory access (DRAM, from CPU)	120 ns	6 min
Solid-state disk I/O (flash memory)	50–150 µs	2–6 days
Rotational disk I/O	1–10 ms	1–12 months
Internet: San Francisco to New York	40 ms	4 years
Internet: San Francisco to United Kingdom	81 ms	8 years
Internet: San Francisco to Australia	183 ms	19 years
TCP packet retransmit	1–3 s	105–317 years
OS virtualization system reboot	4 s	423 years
SCSI command time-out	30 s	3 millennia
Hardware (HW) virtualization system reboot	40 s	4 millennia
Physical system reboot	5 m	32 millennia

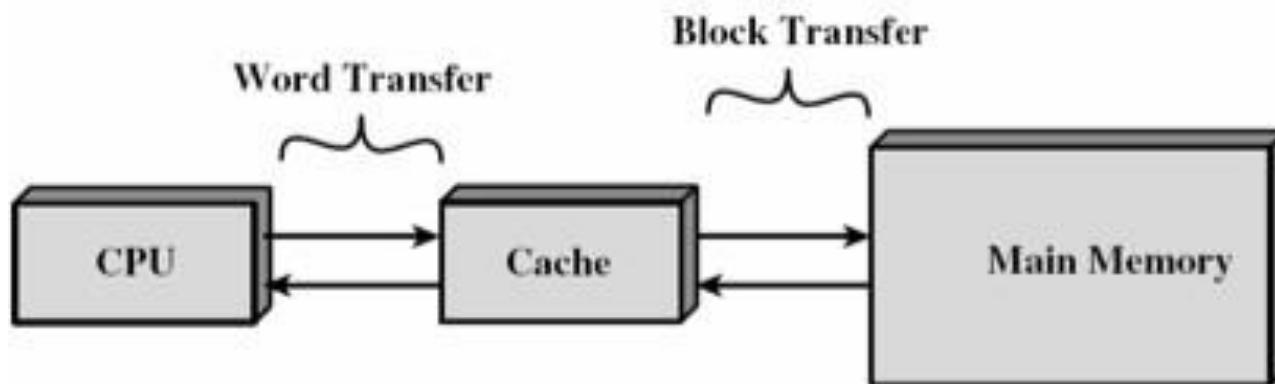
Memoria Caché - Estructura

Se compone de:

- Memoria de datos
- Memoria de etiquetas
- Controlador
 - Selecciona cuantos y cuales bytes se copian a la memoria de datos. Utiliza diferentes algoritmos.

Memoria Caché

- El controlador ve a la RAM en bloques de tamaño fijo
- Por ejemplo de 32 bytes
- Entonces RAM de 1MB son 32768 bloques



Memoria Caché - Ejemplo

- Suponemos RAM de 1 MB ($1M \times 8$).
- Bloques de 32 bytes.
- Caché de 4 K para datos (sin etiquetas).

$$4k = 2^{12} \Rightarrow \frac{2^{12}}{2^5} = 128$$

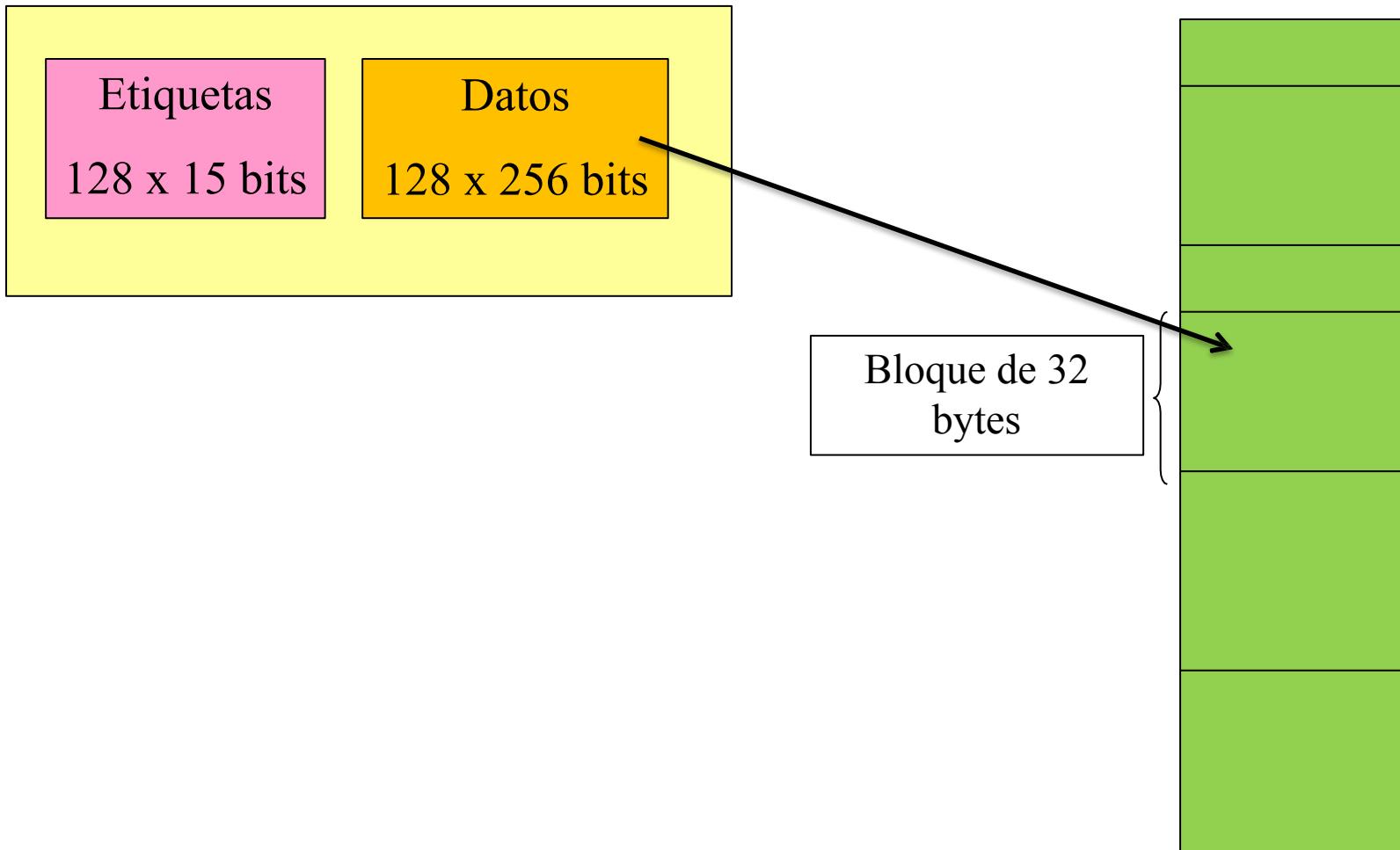
Entonces la memoria de datos tendrá 128 entradas de 256 bits cada una.

Por lo tanto la memoria de etiquetas tendrá 128 entradas de 15 bits cada una porque:

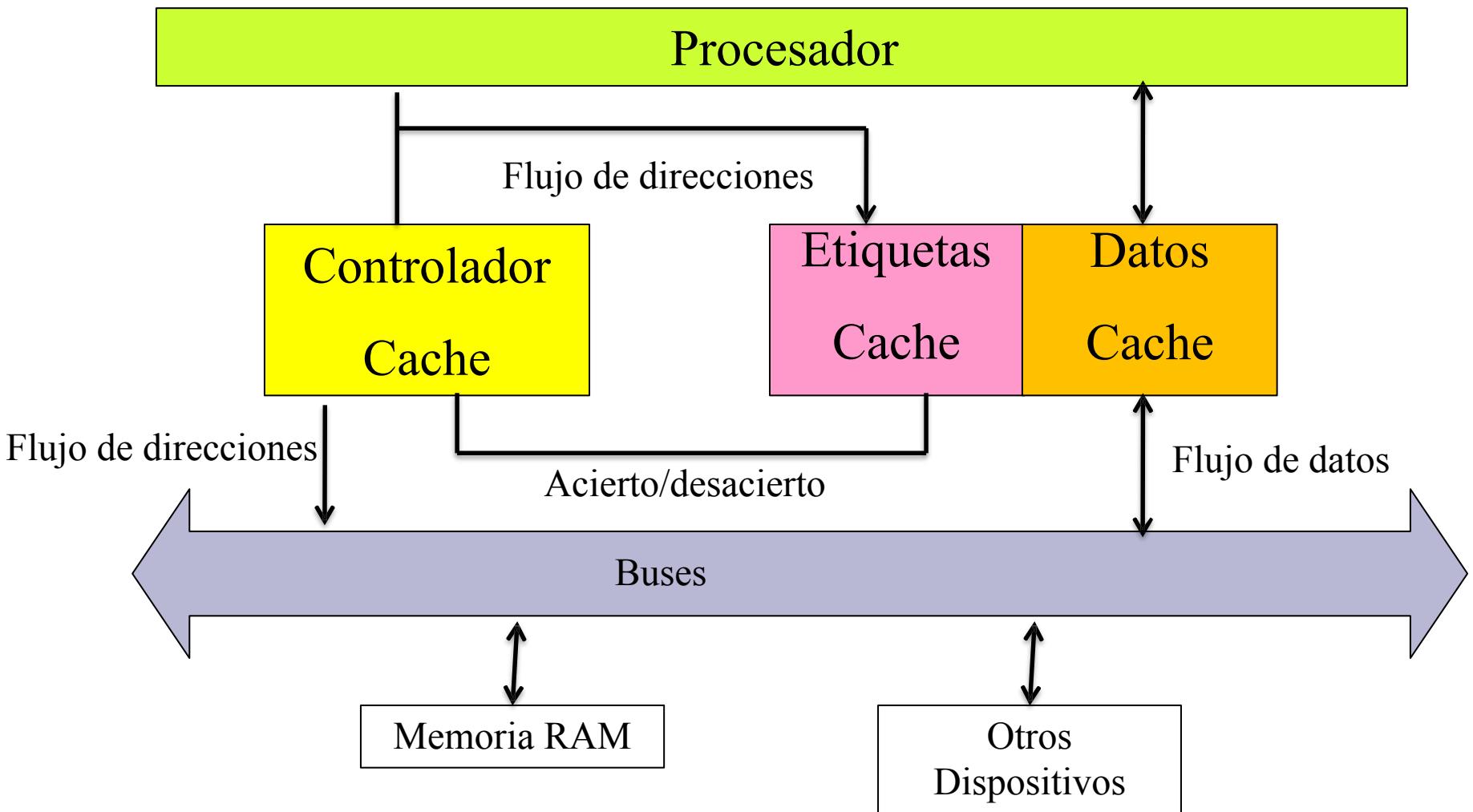
$$\frac{2^{20}}{2^5} = 2^{15}$$

Memoria Caché - Ejemplo

RAM 1M x 8

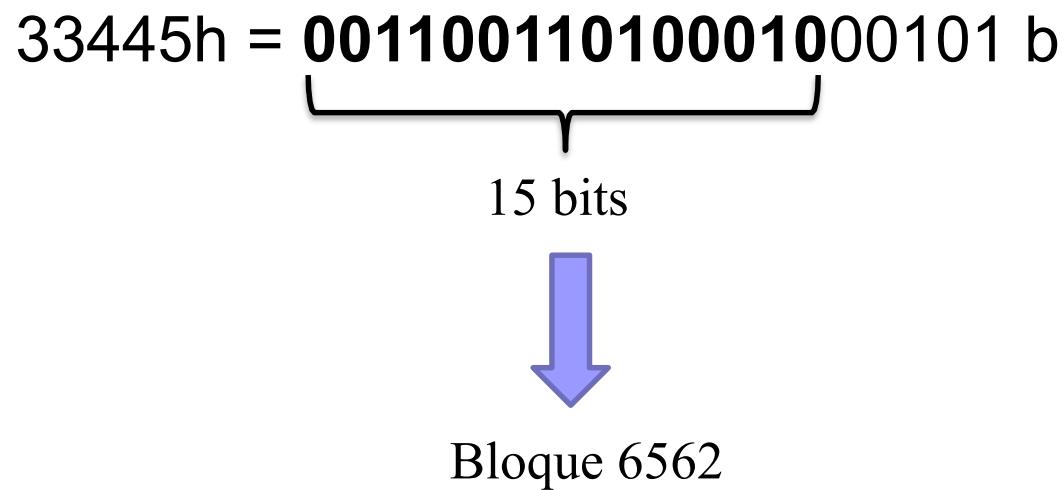


Memoria Caché – Conexión en serie



Memoria Caché - Ejemplo

Suponemos dirección física 33445h



Se busca en las 128 etiquetas si una tiene el contenido
001100110100010

Memoria Caché - Mapeos

- Mapeo Directo.
 - Un bloque de memoria solo se puede mapear a una única ranura en el cache.
 - Sencilla, poco utilizada
- Mapeo Asociativo.
 - Un bloque de memoria se puede mapear a cualquier ranura del cache.
 - Compleja, más utilizada.

Ventajas?

Desventajas?

Memoria Caché – Políticas de sustitución

- Se actualiza la cache al haber fallo o ausencia de palabra buscada.
 - First In First Out (FIFO)
 - Least Recently Used (LRU) (necesita flag !)
 - Random

Memoria Caché – Actualización de RAM

- Escritura inmediata (write through)
 - Se actualizan ambas memorias juntas
 - Más lento
 - Más económico
- Escritura obligada (write back)
 - Solo se actualiza lo estrictamente necesario
 - Más rápido
 - Menos económica
 - Problemas con dispositivos DMA
 - Problemas con Multi-cores