



Sistemas de Procesamiento de Datos – Unidad 6

- Memorias

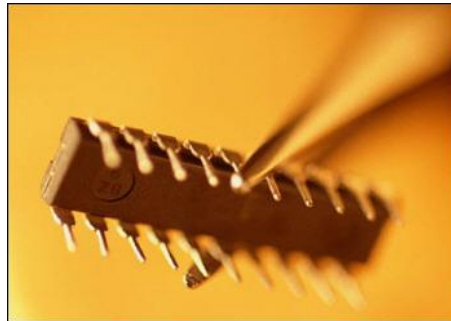
Profesor: Fabio Bruschetti

Ayudante: Pedro Iriso

Ver 2020

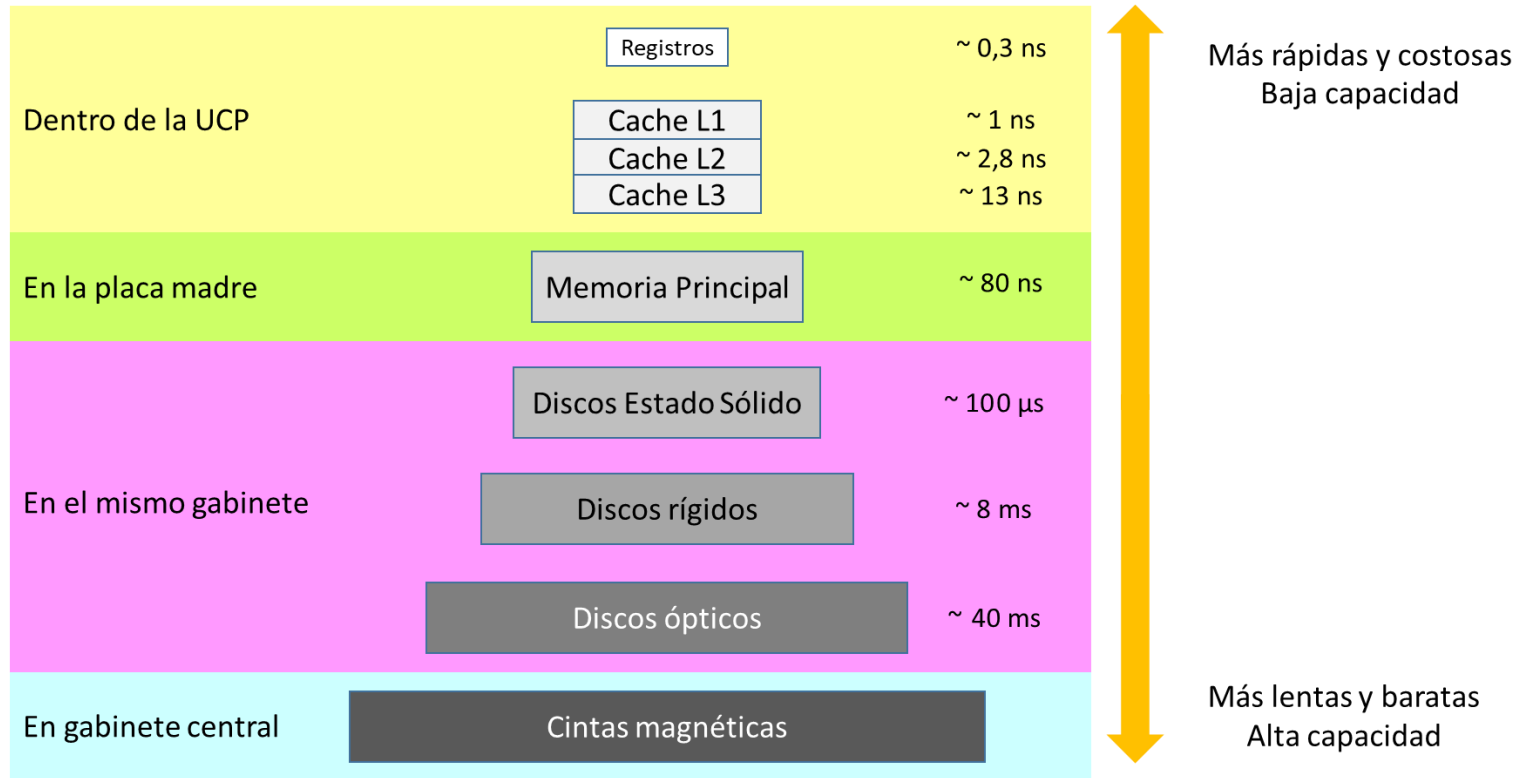
Memoria

- Dispositivo utilizado para almacenar datos
- Características principales
 - Capacidad de almacenamiento
 - Velocidad de acceso (leer, escribir)
 - Permanencia de los datos (volátil, permanente)
 - Presentadas en circuitos integrados (IC's)



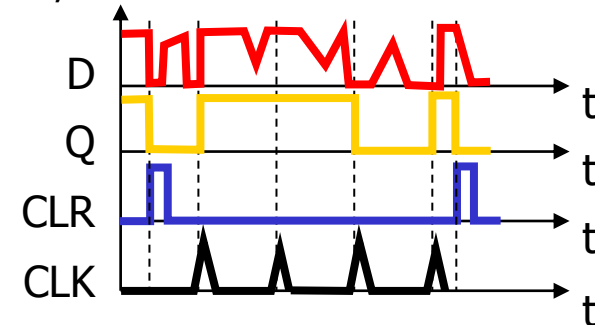
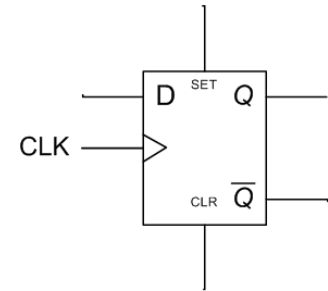
Memoria

■ Jerarquía de Memorias



Memoria – Registros

- Son dispositivos de almacenamiento de n bits o celdas
- Cada celda de memoria almacena un solo bit y está compuesta por un “Flip-Flop” o biestable del tipo “D” que presentan el menor tiempo de acceso
- Un registro de n bits poseerá n flip-flops de este tipo
 - D = Entrada sincrónica
 - Q y \bar{Q} = Salidas
 - CLK = Señal de clock
 - SET = Coloca Q en 1 en forma asincrónica
 - CLR = Coloca Q en 0 en forma asincrónica
- Para guardar un dato (“1” o “0”) en esta celda de memoria, existen dos alternativas:
 - Asincrónicamente:
 - Guardar un “1” \rightarrow SET = 1 y CLR = 0
 - Guardar un “0” \rightarrow SET = 0 y CLR = 1
 - Sincrónicamente
 - Coloco en la entrada D el valor que quiero guardar
 - Coloco un pulso en la entrada CLK



Memoria – Registros

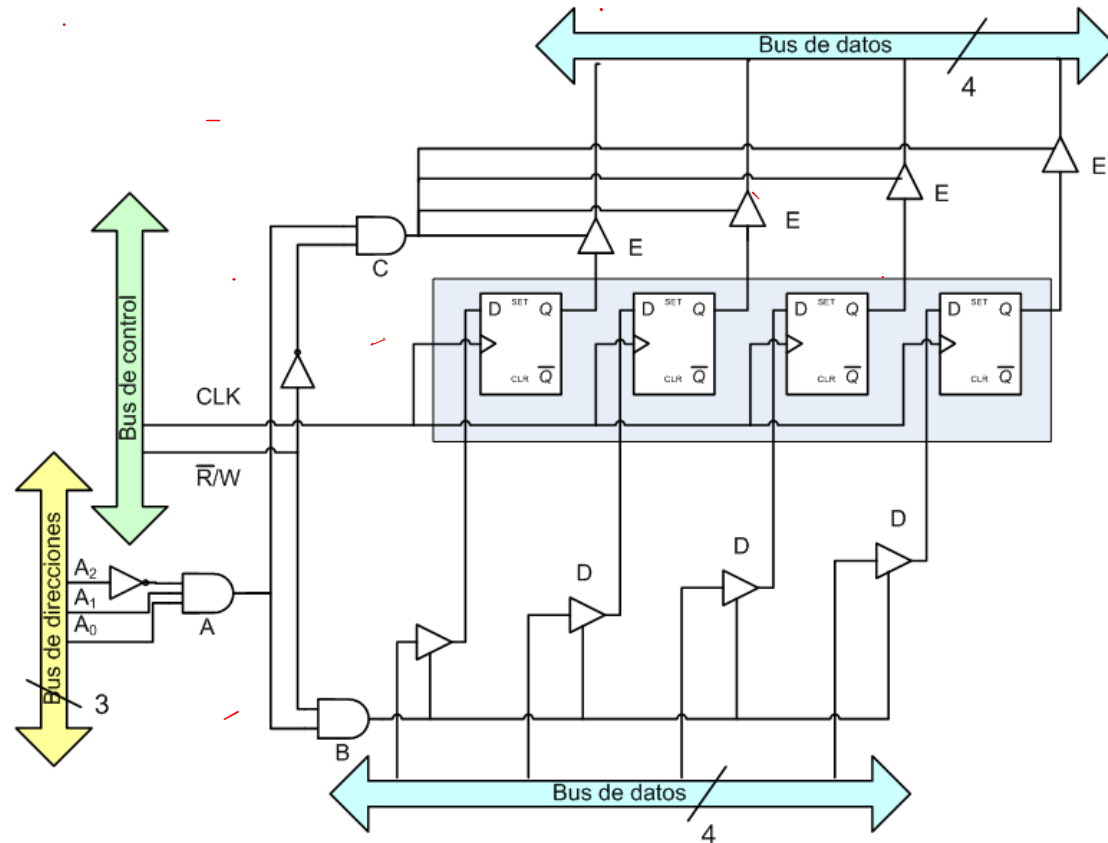
■ Lectura:

- 1) se coloca la dirección "011_b" en el bus de direcciones. La salida de la compuerta "A" es "1"
- 2) La señal R/W se coloca en "0", con lo cual la salida de la compuerta "C" es "1" y habilita los buffers 3-state "E" disponiendo el dato en el bus
- 3) La compuerta "B" tiene un "0" en una de sus entradas lo que impide la escritura del registro

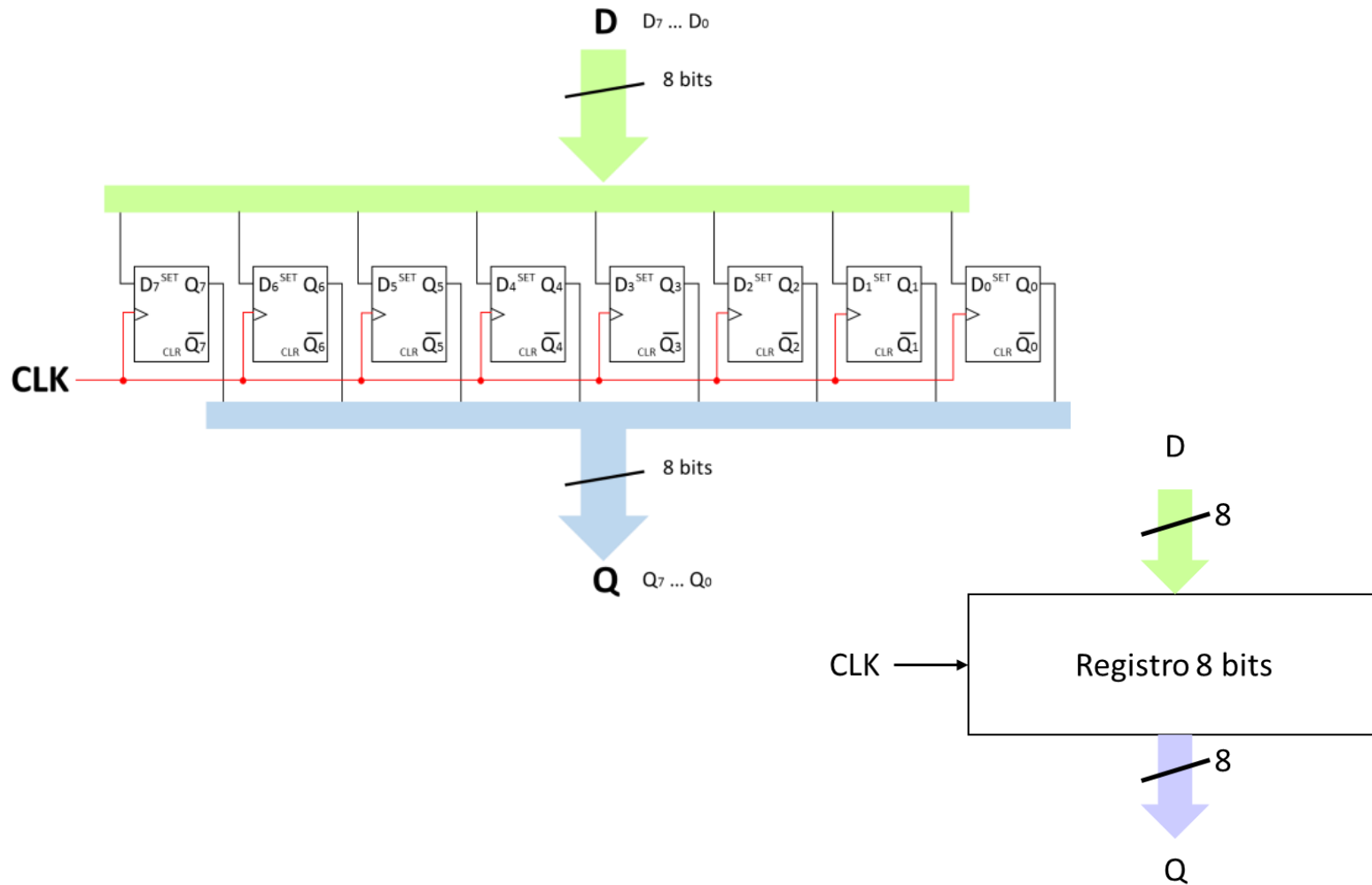
■ Escritura:

- 1) se coloca la dirección "011_b"
- 2) Se presentan los datos a escribir en el bus de datos
- 3) La señal R/W se coloca en "1", con lo cual la salida de la compuerta "B" es "1" y habilita los buffers 3-state "D" disponiendo el dato en cada flip-flop
- 4) La compuerta "C" tiene un "0" en una de sus entradas lo que impide la lectura durante este ciclo

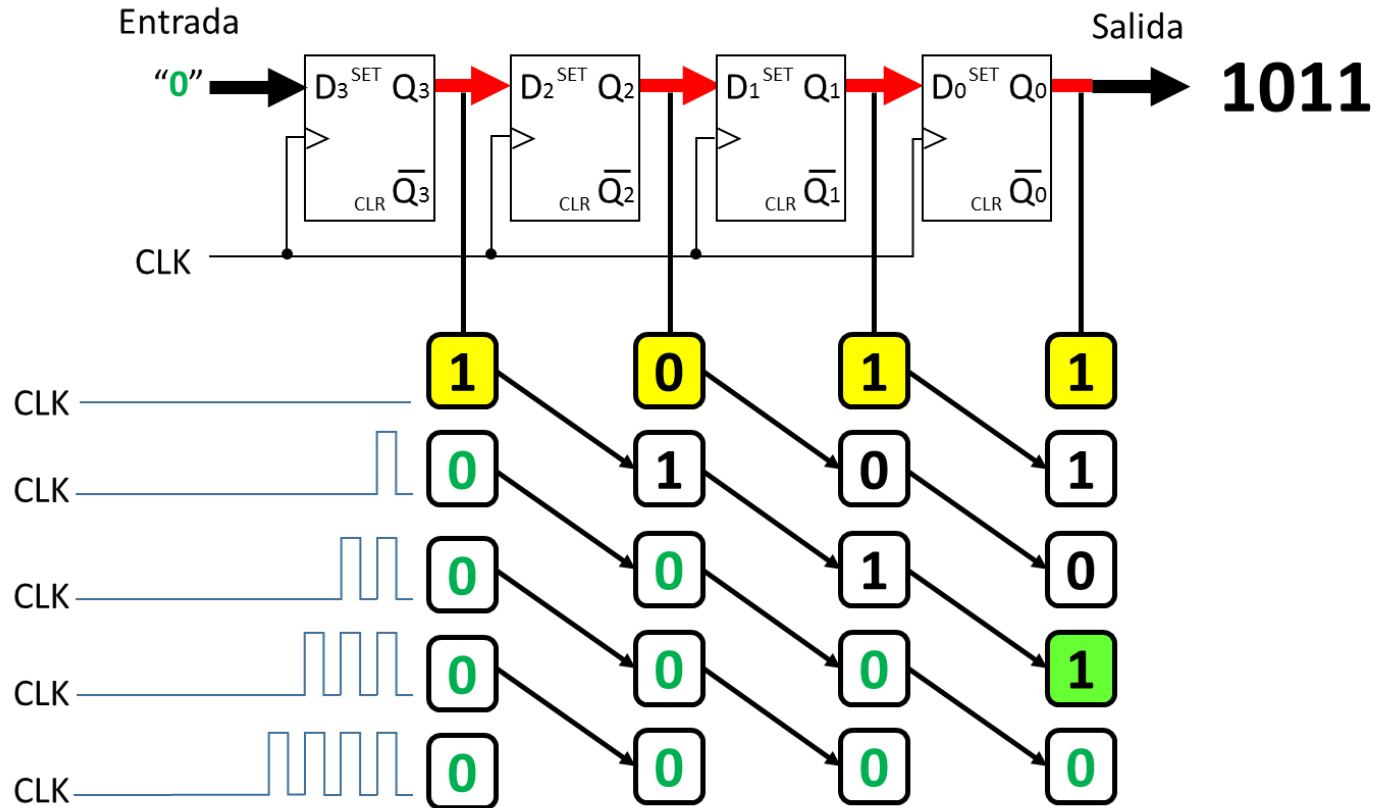
Conexión de un registro de 4 bits a los buses del computador



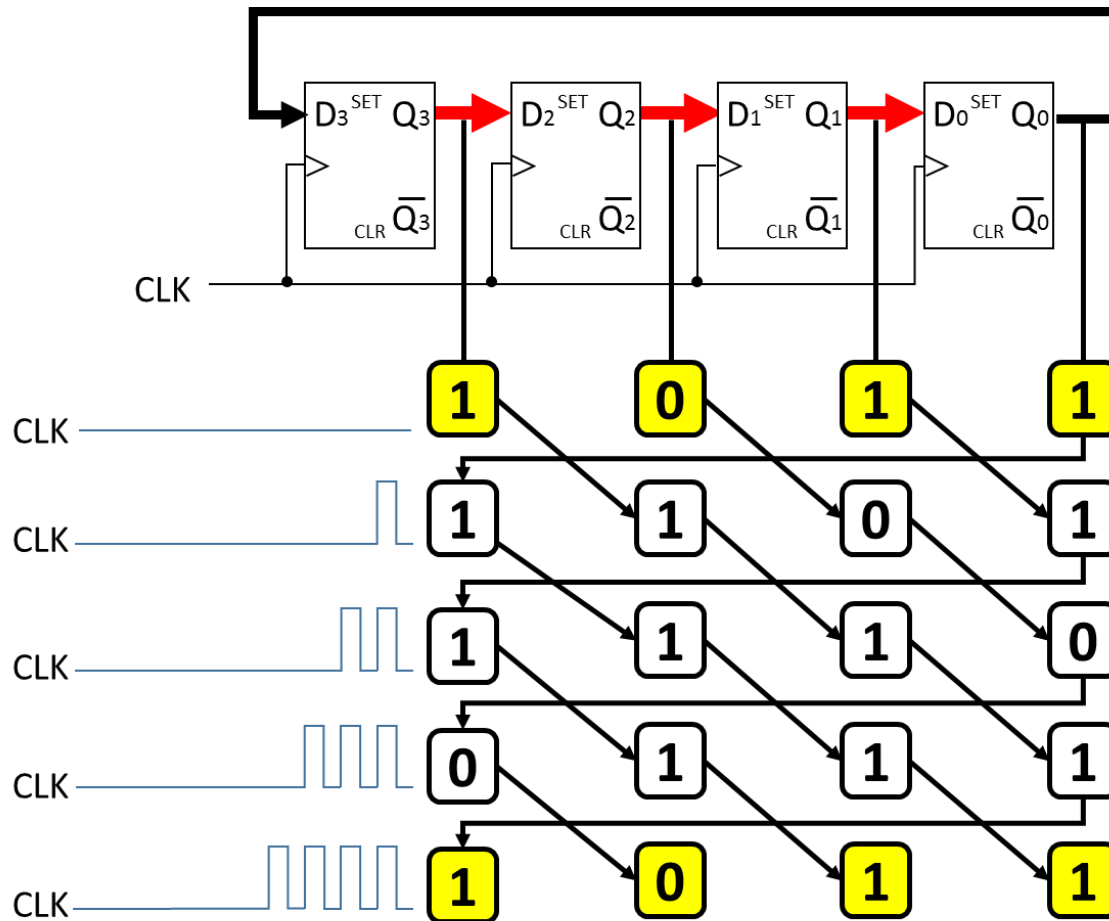
Registros



Memoria – Registros

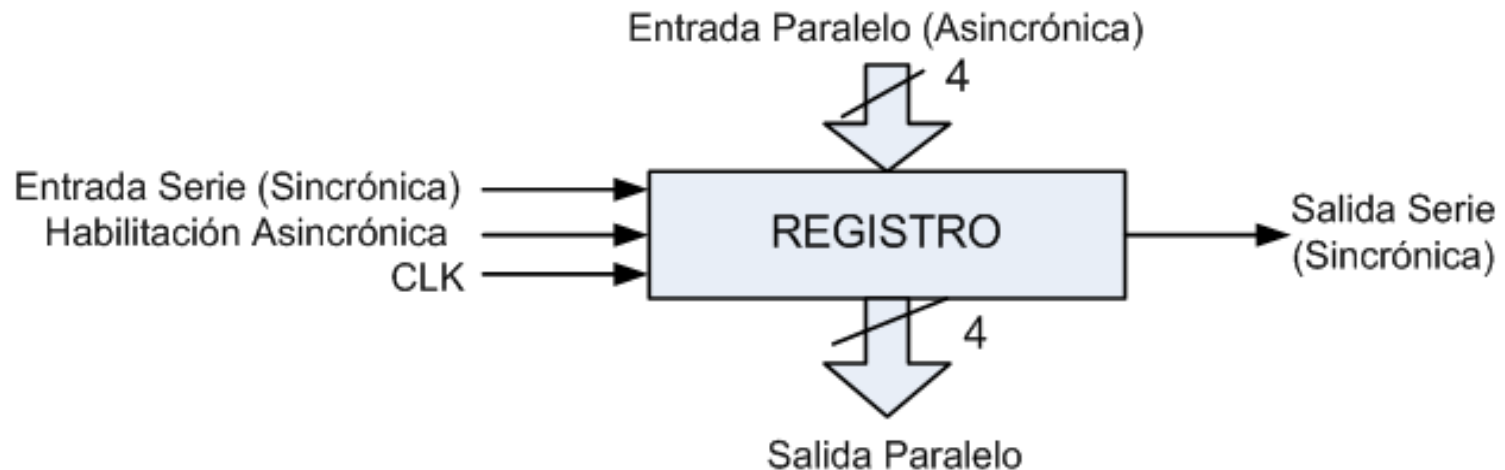


Memoria – Registros



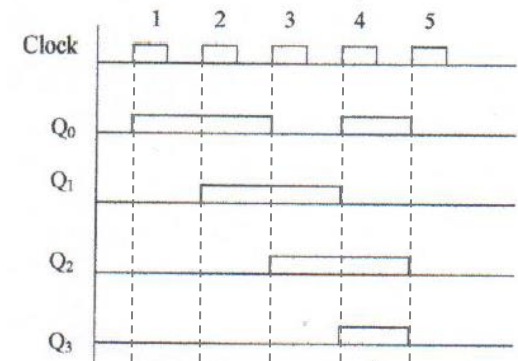
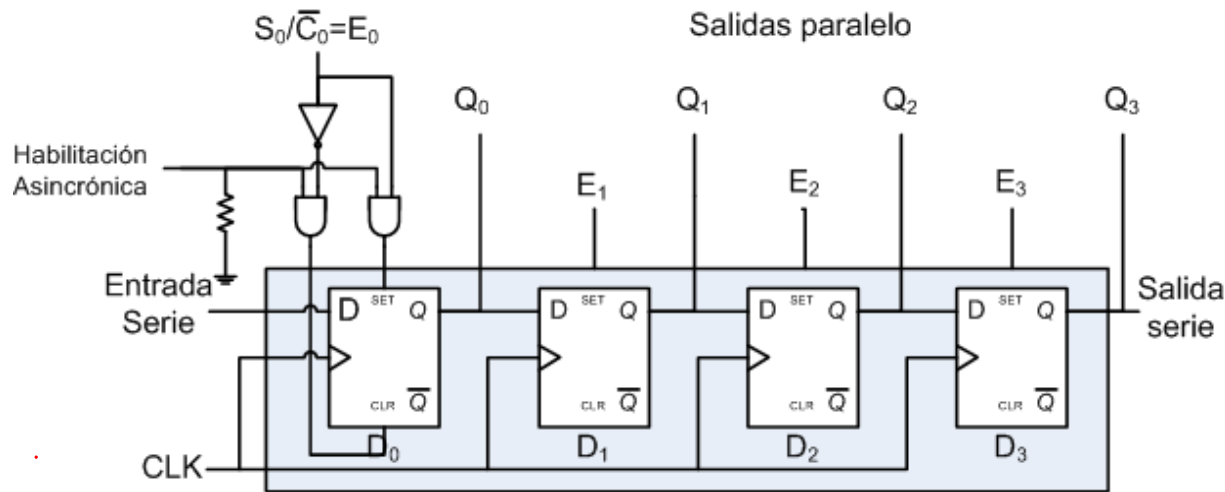
Memoria – Registros

- Registros de desplazamiento
 - Son los que se encuentran en la ALU para hacer las operaciones producto y cociente
 - También se los utiliza para convertir datos serie a paralelo y viceversa
 - Se pueden cargar asincrónica (SET y CLR) o sincrónicamente (con el CLK)
 - Ejemplo de registro de 4 bits



Memoria – Registros

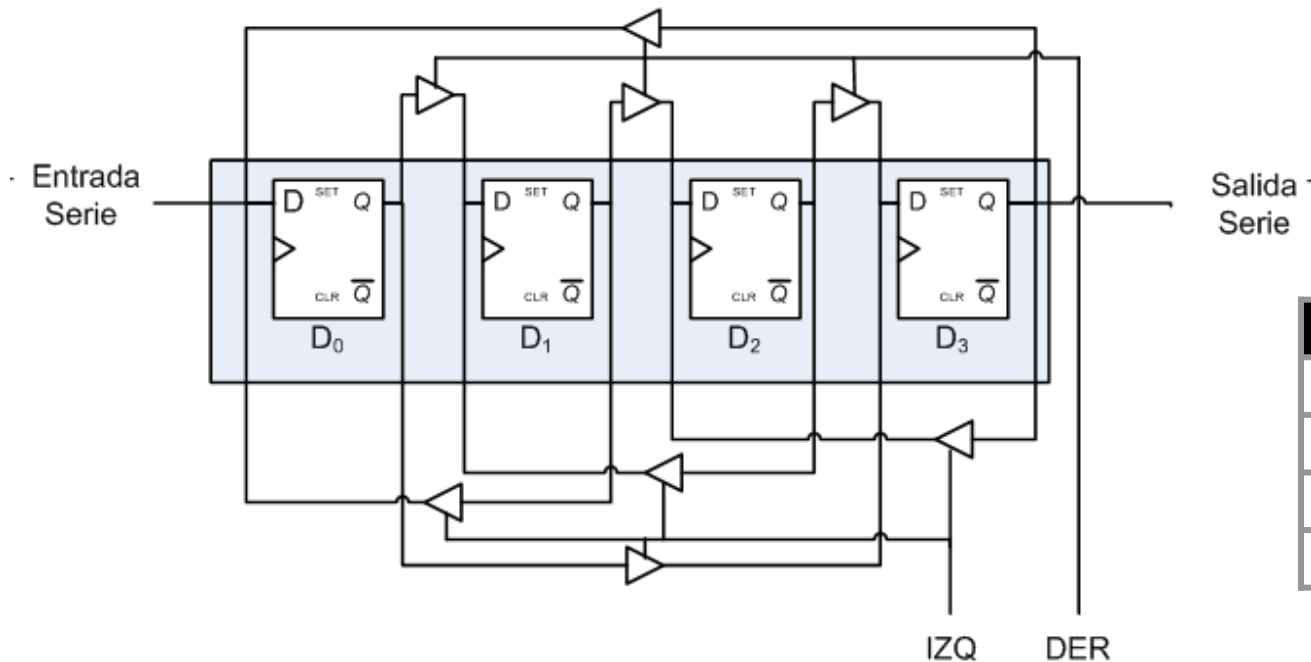
- Registros de desplazamiento
 - Para convertir de serie a paralelo necesito n pulsos de CLK
 - Para convertir de paralelo a serie necesito $n-1$ pulsos de CLK
 - Ejemplo conversión serie \rightarrow paralelo de la combinación "1011"



Memoria – Registros

■ Registros de rotación

- Se cargan asincrónicamente (SET y CLR)
- Luego, los bits circulan de un Flip-Flop a un cualquiera de los dos sentidos, derecha e izquierda
- Se pueden leer los datos en cualquier instante

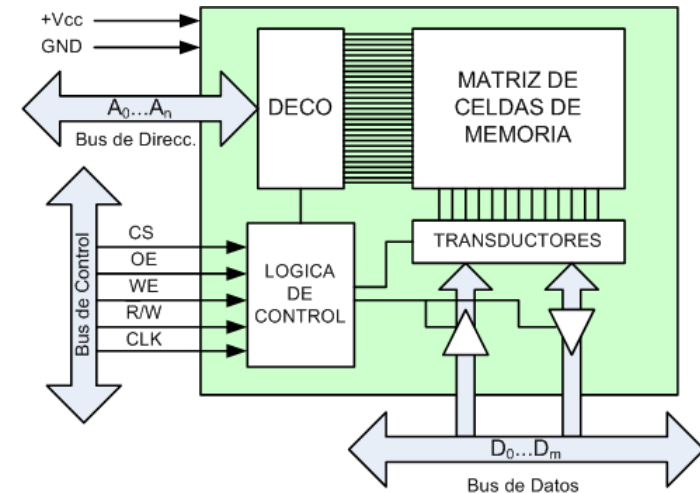
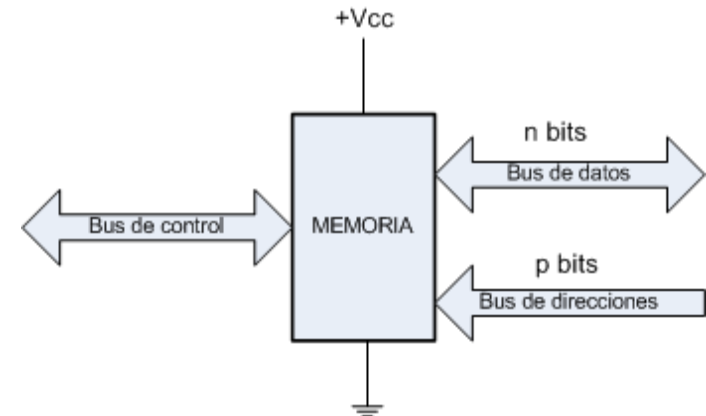


IZQ	DER	Resultado
0	0	Nada
0	1	A derecha
1	0	A izquierda
1	1	Prohibido

Memoria – Principal

■ Memoria Principal

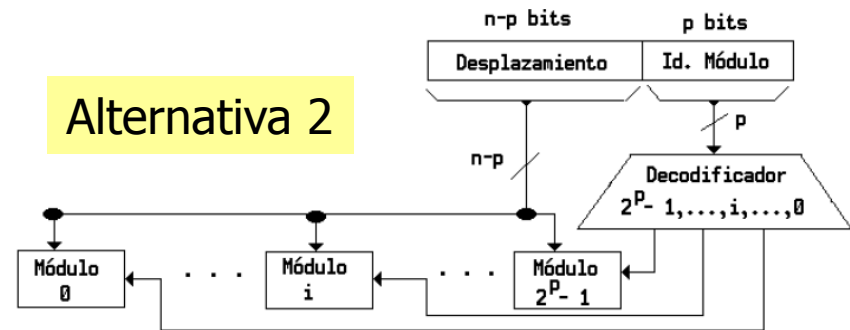
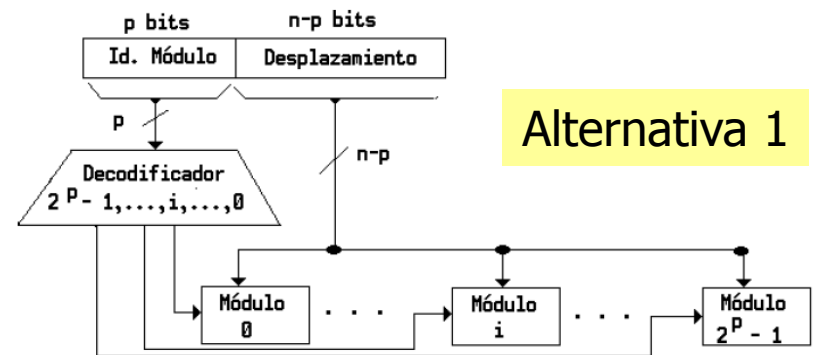
- Dispositivos rápidos conectados y controlados directamente por la CPU
- Suele estar organizada en base a palabras de n bits
- Se accede a cada palabra de a una por vez y a través de una única dirección (de p bits)
- Los datos contenidos en una palabra puede ser una instrucción o un dato que depende únicamente de la interpretación que realice el dispositivo (o proceso) que la requiera
- Existen memorias "interleaved" para accesos a más de una palabra por ciclo (múltiples accesos)
- RAS y CAS = Acceso



Memoria – Interleave

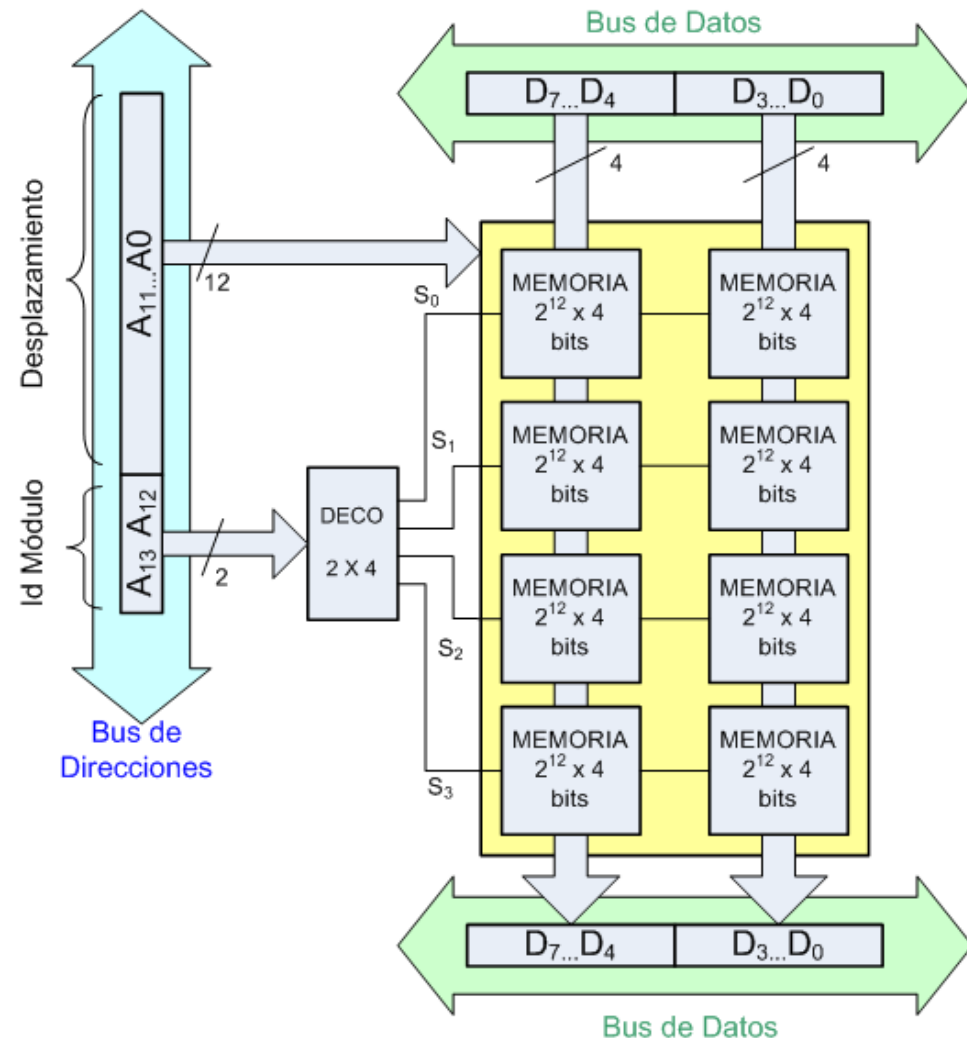
Iterleave – Intercalación de Direcciones

- 2^p módulos de $2^{(n-p)}$ direcciones de memoria
- En cada módulo guardo $2^{(n-p)}$ direcciones de memoria
- Alternativa 1: consecutivas
- Alternativa 2: separadas 2^p



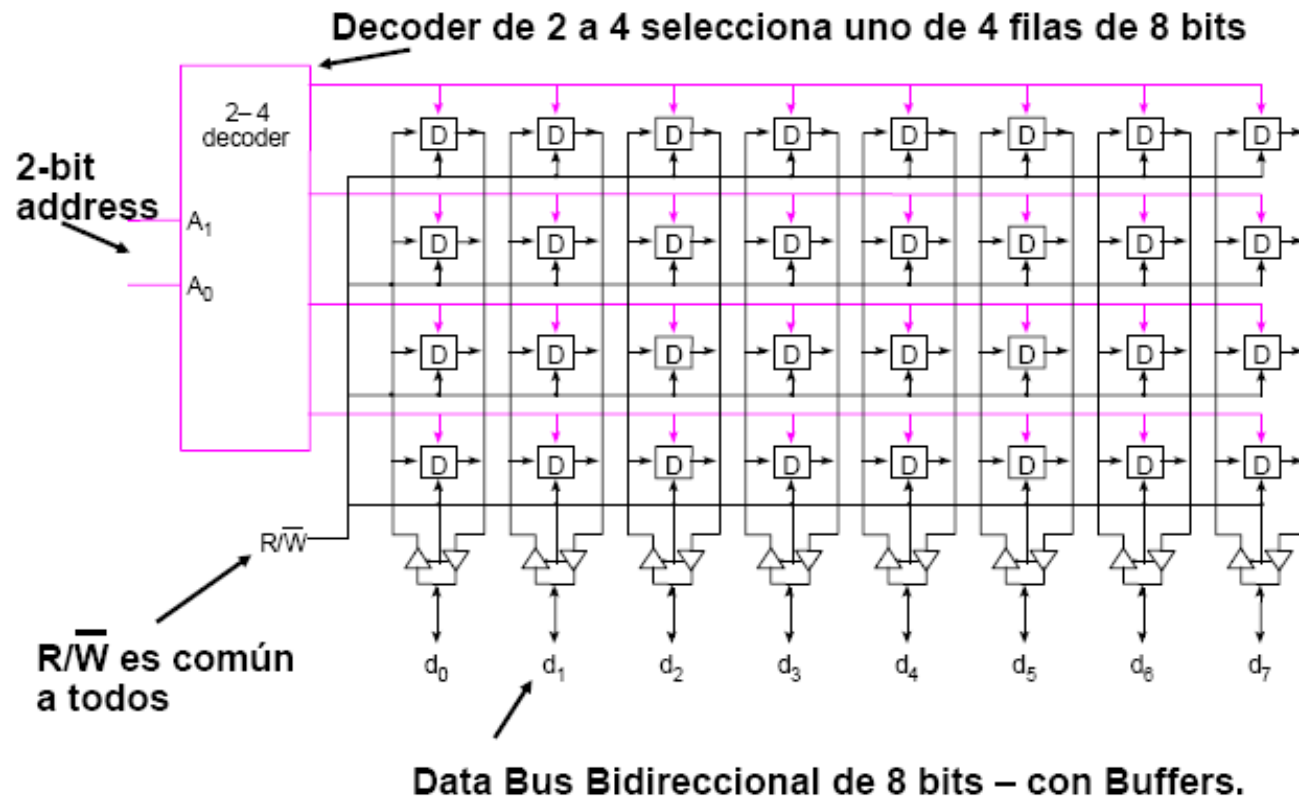
Memoria – Bancos

- Bancos de memoria
 - La Memoria Principal se agrupa físicamente en varios módulos (Bancos)
 - En un banco, la palabra direccionada podrá obtenerse:
 - Como la unión de varias partes de la palabras situadas cada una en módulos diferentes → encolumnar más de un módulo por dirección
 - En módulos distintos dependiendo de la dirección requerida → Más de una fila de módulos



Memoria – Bancos

- Ejemplo Memoria 2D de 4 x 8



Memoria – Ordenamiento

- Ordenamiento de las palabras, byte por byte

- Palabra a ordenar es

- 9FCD = 1001111111001101

→ Big Endian
← Little Endian

- Little Endian (Intel)

- 1100110110011111

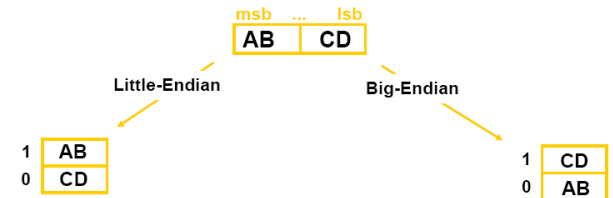
- Primero el LSByte, luego el MSByte

- Big Endian (Motorola)

- 1001111111001101

- Primero el MSByte, luego el LSByte

Ejemplo: El número Hexa de dos bytes \$ABCD, a guardar en el lugar 0:



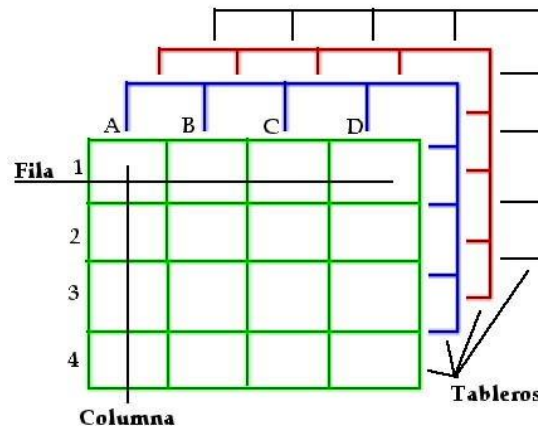


Memorias – Tiempos

- Características / Tiempos
 - Access Time (Tiempo de Acceso)
 - Desde que se realiza un pedido de lectura hasta que el mismo queda satisfecho.
 - Cycle Time (Tiempo de Ciclo)
 - Desde que se realiza un pedido hasta que se puede realizar el siguiente.
 - Write Time (Tiempo de Escritura)
 - Al escribir un dato, tanto el dato y la dirección donde se escribirá deben estar presentes antes que llegue la señal de escritura (Setup Time)
 - Y deben mantenerse luego que se retire la señal de escritura (Hold Time)
 - Los tiempos de acceso y de ciclo son muy importantes. Un CPU constantemente accede a la Memoria, y esto representa una traba a su velocidad.

Memorias – Latencias

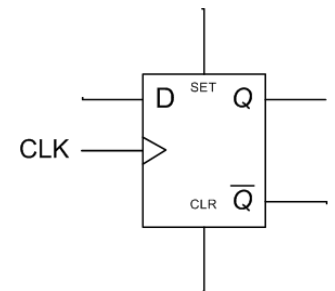
- Las más importantes son:
 - CAS (Column Access Strobe): indica el tiempo que tarda la memoria en colocarse sobre una columna.
 - RAS (Row Access Strobe): indica el tiempo que tarda la memoria en colocarse sobre una fila.
 - ACTIVE: indica el tiempo que tarda la memoria en activar un tablero.
 - PRECHARGE: indica el tiempo que tarda la memoria en desactivar un tablero.



Memorias – Tipos

- Memorias SRAM

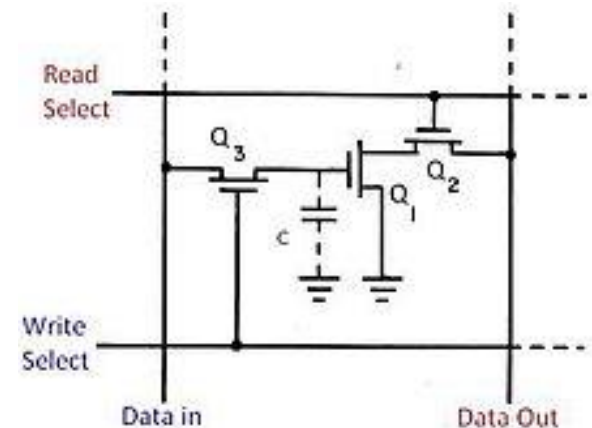
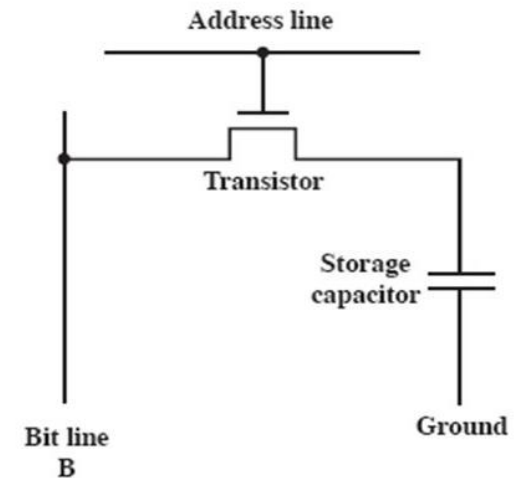
- Static RAM: Memoria de lectura/escritura que mantiene sus datos siempre que tenga alimentación.
- Características
 - Son muy rápidas (devuelven el dato en 1 ns)
 - Muy caras
 - Consume mucha energía
- Usadas en dispositivos que requieren alta velocidad de operación
 - Memoria caché de microprocesadores
 - Procesamiento digital de imágenes



Memorias – Tipos

■ Memorias DRAM

- Dynamic RAM: A diferencia de las SRAM, necesitan un pulso de energía periódico (refresco) para poder mantener los datos.
- Características
 - Más lentas que las SRAM (access time = 80 ns)
 - Más baratas que las SRAM
 - Consumo de energía mucho mas bajo
- Usadas como Memoria Principal en todo sistema con microprocesadores.
- Se convirtieron en el cuello de botella en un sistema moderno (PC).



Memorias – Tipos

■ Memorias SDRAM

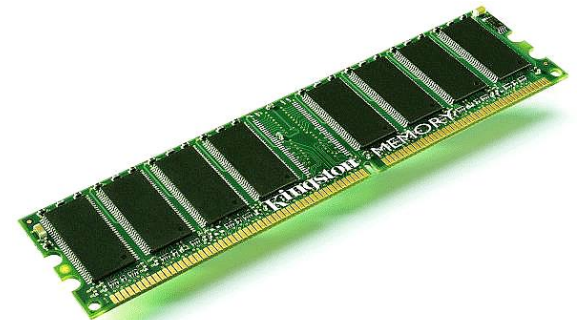
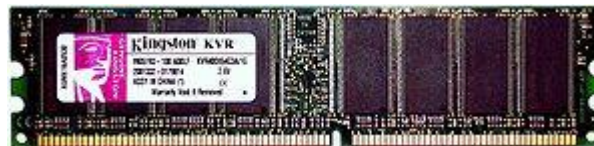
- Synchronous DRAM: Permite el acceso a un bloque de datos que estén en fila, haciendo la transferencia sincronizada.
- Características
 - Permite transferencias en ráfagas (burst)
 - Se envía primero cantidad de datos a transferir
 - Luego la dirección donde se comenzará a almacenar
 - A partir de allí se transfieren varios bytes por ciclo.
 - Incorporan un circuito para hacer el refresco automáticamente haciendo que su velocidad de transferencia de datos sea más alta.
- Muy comunes en las PCs hasta hace un par de años.



Memorias – Tipos

■ Memorias DDR DRAM

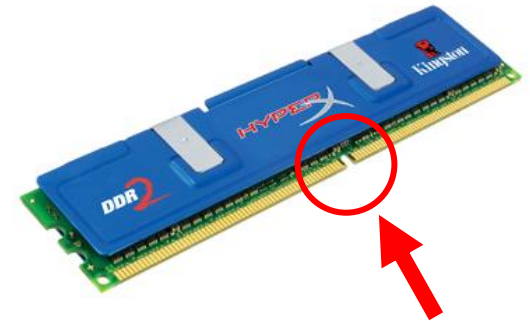
- Double Data Rate: Fabricadas con tecnología SDRAM pero transfieren los datos en ambos flancos del reloj, duplicando la tasa de transferencia
- Trabajan con 2.5V en lugar de los 3.3V con que trabajan las SDRAM
- Adoptadas inicialmente por sistemas equipados con AMD mientras que Intel utilizaba RAMBUS
- Características
 - Poseen además técnicas avanzadas de optimización, que las hacen más veloces
 - Interleaving
 - Pipelining
 - El tiempo de acceso no es siempre igual, el primero dura más.
- Buffer interno de 2 bits (prefetch buffer)
- Velocidades del buffer desde 200 Mhz hasta 400 Mhz
- DIMM de 184 contactos



Memorias – Tipos

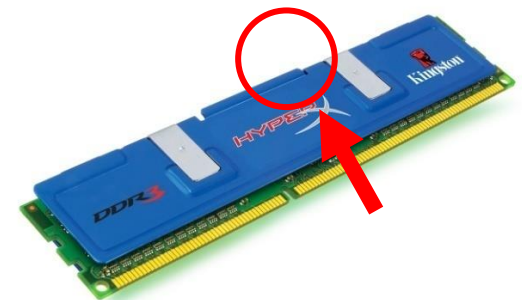
■ Memorias DDR2

- Duplican la velocidad del bus respecto de las DDR (533 MHz a 1 GHz)
- El buffer de prefetch es de 4 bits
- Tensión de trabajo de 1.8V
- DIMM de 240 contactos
- A pesar de todo, tienen menor latencia que las DDR



■ Memorias DDR3

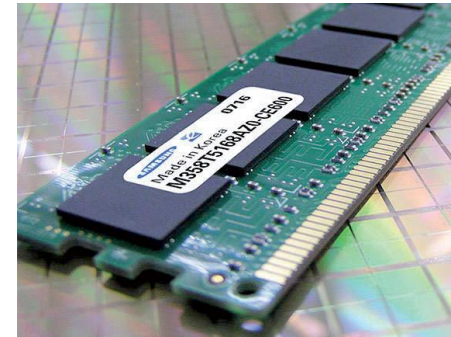
- El buffer de prefetch es de 8 bits
- Velocidad del bus de 800 Mhz a 2 Ghz
- Tensión de trabajo de 1.5V
- También el módulo DIMM es de 240 contactos pero se modifica la muesca
- Tienen menor latencia que las DDR2!



Memorias – Tipos

■ Memorias DDR4

- Duplican la velocidad del bus respecto de las DDR (2,13 GHz a 4,23 GHz)
- El buffer de prefetch es de 4 bits
- Tensión de trabajo de 1,2 a 1,05 V
- DIMM de 288 contactos
- No es compatible con las anteriores
- Lanzada a finales de enero de 2014
- Incremento del largo de la ráfaga de datos
- Chequeo de paridad a nivel de bus de address y comandos



■ Memorias DDR5

- Velocidad del bus 5,2 GHz o más
- Tensión de trabajo de 1,1 V
- 128 GB o más en una placa
- Duplicarán ancho de banda de las DDR4
- Se comercializará cerca de julio 2020



Memorias – Tipos



- XDR (Extreme Data Rate)

- Una versión mejorada de la memoria RAMBUS RDRAM (usada en Nintendo 64)
- Permite un ancho de banda mucho más alto que las DDR2 u GDDR4
- Ideal para placas de video o consolas de videojuegos (PlayStation 3)

- MRAM (Magnetoresistive RAM)

- Usa discos ferromagnéticos separados por una capa aislante (en lugar de acumulación de cargas eléctricas) para guardar un bit
- Ventajas: Menores tiempos de acceso y menor consumo de potencia
- Desventajas: Menor densidad (180 nm) de integración
- No volátil !!!! (Booteo instantáneo)

- Otras Características

- ECC / NonECC
- Buffered / Unbuffered





Memorias – Clasificación

Clasificación por	Tipos
Método de acceso	Aleatorias Semialeatorias Secuenciales
Velocidad / Tiempo de acceso	Memoria interna del proc. (cache) Memoria Principal Memoria Secundaria
Según la forma de ubicar la información	Acceso por dirección - RAM Acceso por contenido - CAM
Según el espacio de direccionamiento	Memoria Real Memoria Virtual
Según la capacidad de modificación de la información almacenada	RAM ROM PROM EPROM
Según la perdurabilidad del dato almacenado	DRO (Destructive Read Out) / NDRO Dynamic Storage / Static Volátiles / No volátiles (discos)



Memorias – Clasificación

- Según el método de acceso
 - Aleatorias (**Random Access Memory**)
 - Permiten direccionar cualquier posición de la memoria de forma directa e independiente del lugar en donde se encuentre
 - Semi-aleatorias
 - Discos floppy, discos duros, CD-ROM
 - Secuenciales
 - Para acceder a una posición de memoria, debo recorrer la memoria desde el principio
 - Cintas magnéticas



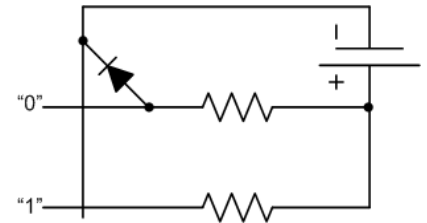
Memorias – Clasificación

- Según su volatilidad
 - Volátiles
 - Al desconectarlas del suministro eléctrico, se pierden los datos
 - Estáticas
 - Construidas con flip-flops, y éstos con transistores
 - No necesita refresco
 - Consumen más energía
 - Más veloces
 - Dinámicas
 - Construidas sencillamente con capacitores
 - Alta densidad de integración
 - Necesitan refresco
 - No volátiles
 - Los datos permanecen aún sin ser energizadas
 - ROM
 - PROM
 - EPROM
 - EEPROM

Memorias – Clasificación

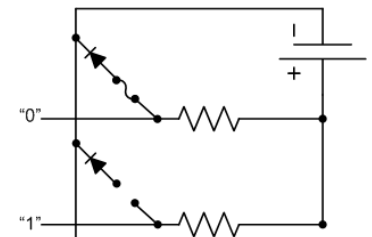
- ROM (**R**ead **O**nly **M**emory)

- Se graban los datos durante el proceso de fabricación y no se puede alterar
- Se coloca un diodo en donde se quieren guardar "ceros"



- PROM (**P**rogrammable **R**ead **O**nly **M**emory)

- El fabricante coloca un diodo en cada celda de memoria, junto con un fusible en serie. De esta manera, en todas las celdas hay guardado un "cero"
- En donde se necesita guardar un "uno", hay que quemar el fusible. Esto se realiza mediante un programa externo
- Se programa una sola vez



Memorias – Clasificación

- EPROM (Electrically Programable Read Only Memory)
 - Las programa el usuario en forma eléctrica
 - Se borra mediante luz ultravioleta
- EEPROM o E²PROM (Electrically Erasable Programmable Read Only Memory)
 - También las programa el usuario en forma eléctrica
 - Se borran de a celda por celda con una tensión eléctrica de polarización invertida



Memorias – Clasificación

■ FLASH

- Memoria del tipo EEPROM
- Se pueden escribir y borrar varias celdas simultáneamente
- Fabricadas con compuertas NOR y NAND para cada celda
- Barata, rápida (hasta 20 MB/s), de bajo consumo
- Durables: escribir y borrar su contenido una vez por día durante 27 años (Toshiba)





Memorias – Clasificación

■ Sus Usos

Type	Category	Erasure	Byte alterable	Volatile	Typical use
SRAM	Read/write	Electrical	Yes	Yes	Level 2 cache
DRAM	Read/write	Electrical	Yes	Yes	Main memory
ROM	Read-only	Not possible	No	No	Large volume appliances
PROM	Read-only	Not possible	No	No	Small volume equipment
EPROM	Read-mostly	UV light	No	No	Device prototyping
EEPROM	Read-mostly	Electrical	Yes	No	Device prototyping
Flash	Read/write	Electrical	No	No	Film for digital camera

Memoria Cache

- Es una cantidad limitada de memoria de alta velocidad insertada entre el procesador y la memoria principal para mantener una porción del contenido de la memoria principal que está actualmente en uso
- El objetivo es incrementar la performance del sistema CPU \leftrightarrow Memoria

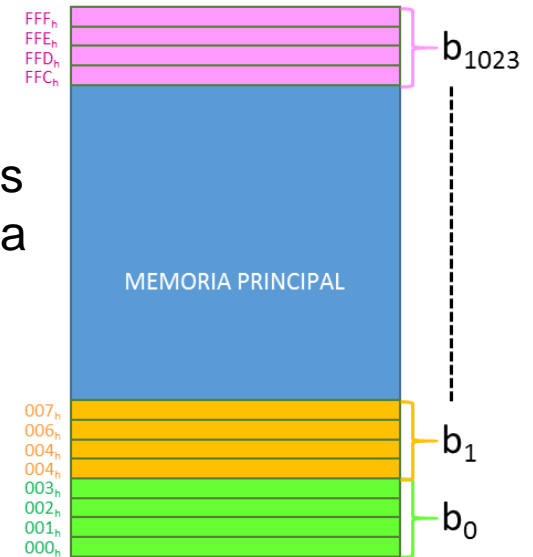


- Cuando se requiere acceder a una dirección de memoria, con altísima probabilidad se requerirá acceder a la siguiente \rightarrow Se “cachean” bloques contiguos de memoria
- Su éxito puede ser atribuido a la propiedad de localidad de referencia
- Acierto (Hit): El procesador encuentra en la memoria Cache el dato que busca
- Falla (Miss): El procesador NO encuentra en la memoria Cache el dato que busca, debe buscarla en la memoria RAM común

Memoria Cache

■ Organización

- Indica las reglas para copiar datos de la memoria principal a la memoria Cache, como así también las reglas para aloca datos nuevos cuando la memoria Cache está llena
- Se asumirá:
 - La memoria común tiene una capacidad de 2^m bytes
 - Se la divide en bloques consecutivos de b palabras
 - El tamaño del bloque es una potencia de 2
 - Se tendrá la cantidad de $2^m / b$ bloques
- Existirán tres escenarios
 - Asociativa (Full Associative)
 - Mapeo Directo (Direct Mapping)
 - Asociativa Agrupada (Set Associative) solución de compromiso entre las dos anteriores



- $000_h - FFF_h$
- 12 bits $\rightarrow 2^{12}$
- 4096 direcciones
- $B = 4$
- $4096/4 = 1024$ bloques

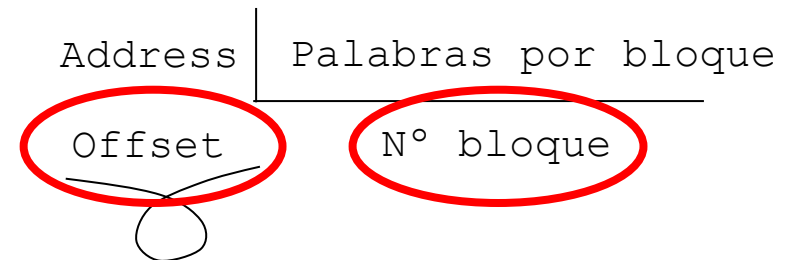
Organización asociativa

- M cantidad de líneas o renglones
- Cada renglón tiene
 - B = Bit de validez
 - N = #bloque
 - V = Valores

Memoria Caché

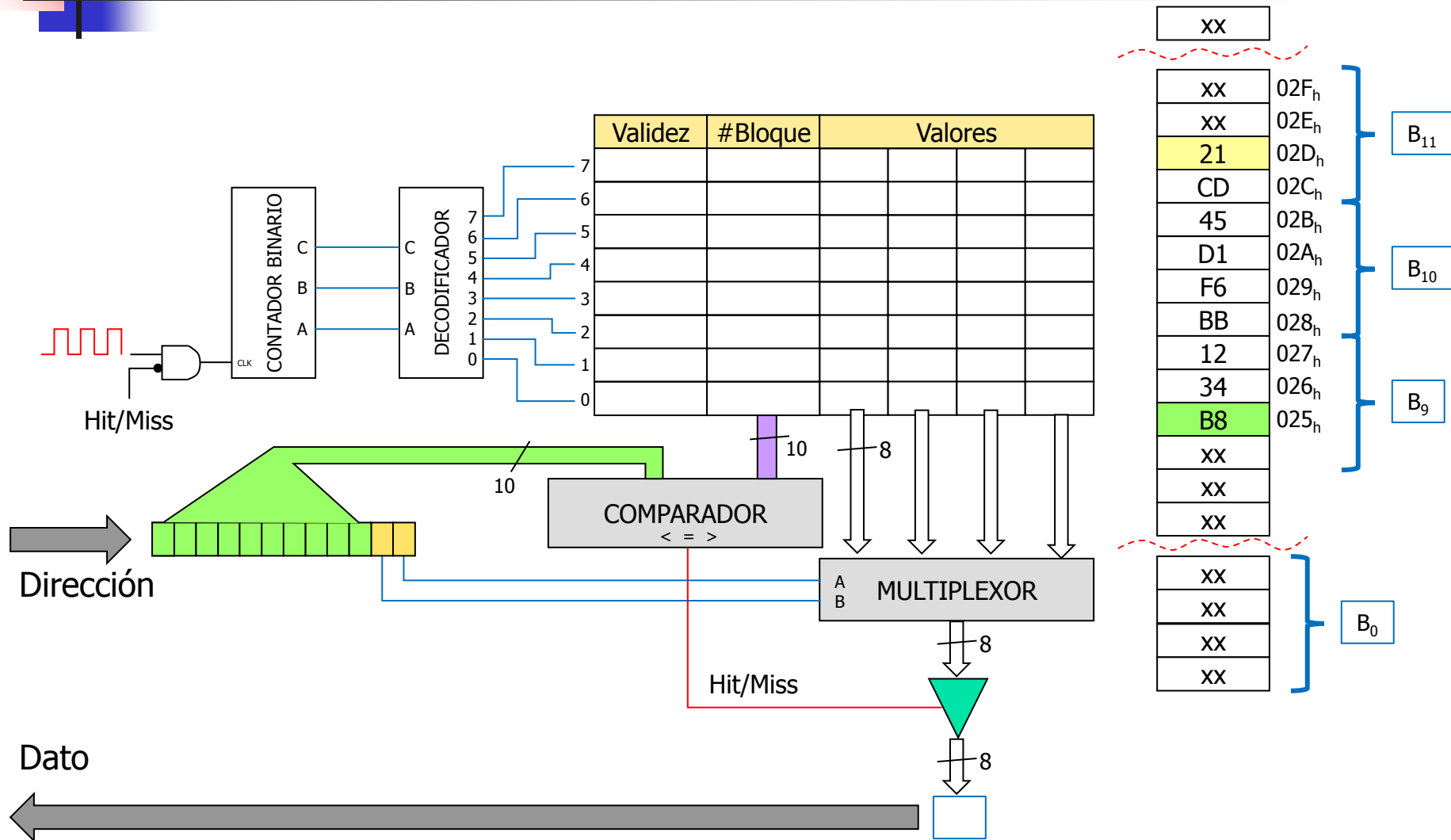
Renglón	V	N	V			
7						
6						
5						
4						
3						
2						
1						
0						

- Los datos de cualquier dirección de memoria común pueden ser almacenados en cualquier dirección de la memoria Cache
- El microcódigo debe
 - calcular el número de bloque
 - buscar ese número de bloque en la Cache
 - Circuitería adicional
 - Mas costosa
- Se necesitan sofisticados algoritmos de búsqueda
- Ofrece la mejor proporción teórica de aciertos

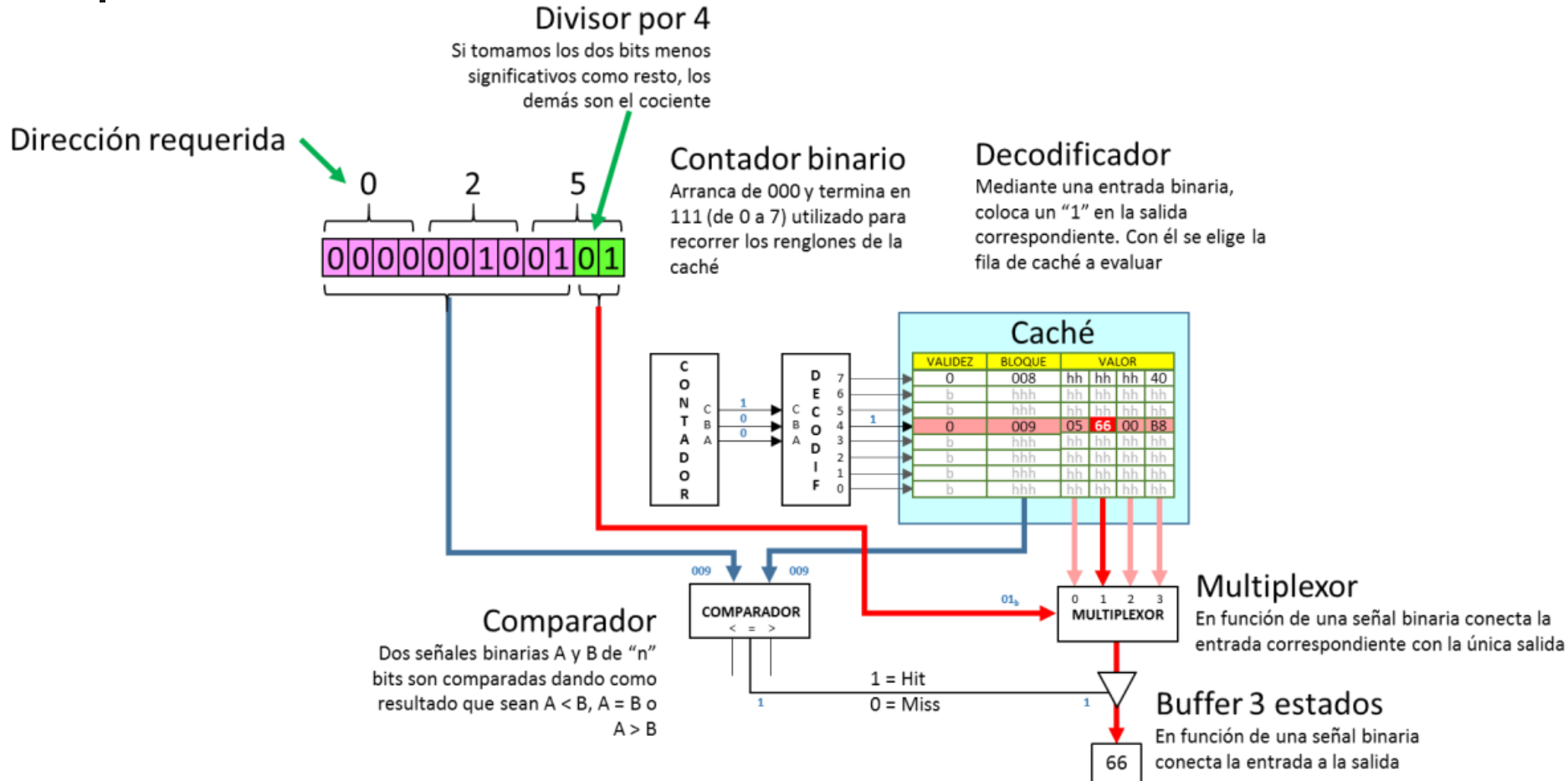


Organización asociativa

Address | Palabras por bloque
 Offset | N° bloque



Organización asociativa



Organización mapeo directo

- M cantidad de líneas o renglones

- b = palabras por bloque

- Cada renglón tiene

- B = Bit de validez
- E = Etiqueta
- V = Valores

Memoria Caché

Renglón	B	E	V			
7						
6						
5						
4						
3						
2						
1						
0						

Nros. De bloque

7	15	23	31
6	14	22	30
5	13	21	29
4	12	20	28
3	11	19	27
2	10	18	26
1	9	17	25
0	8	16	24

Etiqueta → 0 1 2 3

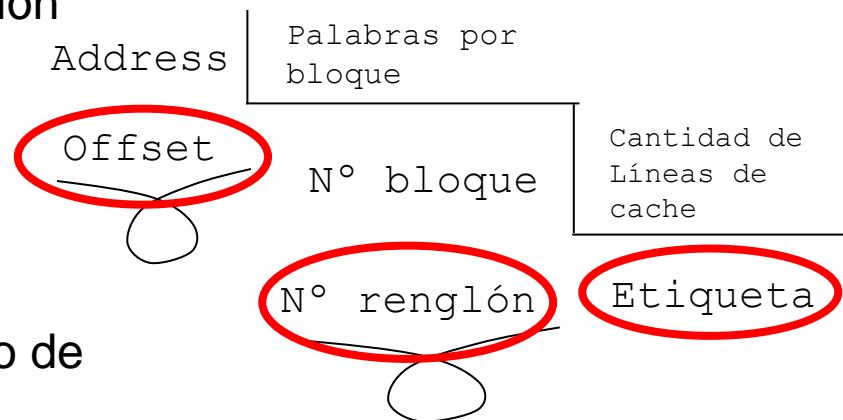
- Cada bloque tiene asignado un único renglón

- Nro. de bloque = dirección / b
- Etiqueta = $\# \text{bloque} / \text{Cant. reng}$
- Nro. de renglón = $\# \text{bloque} \bmod \text{Cant. reng}$

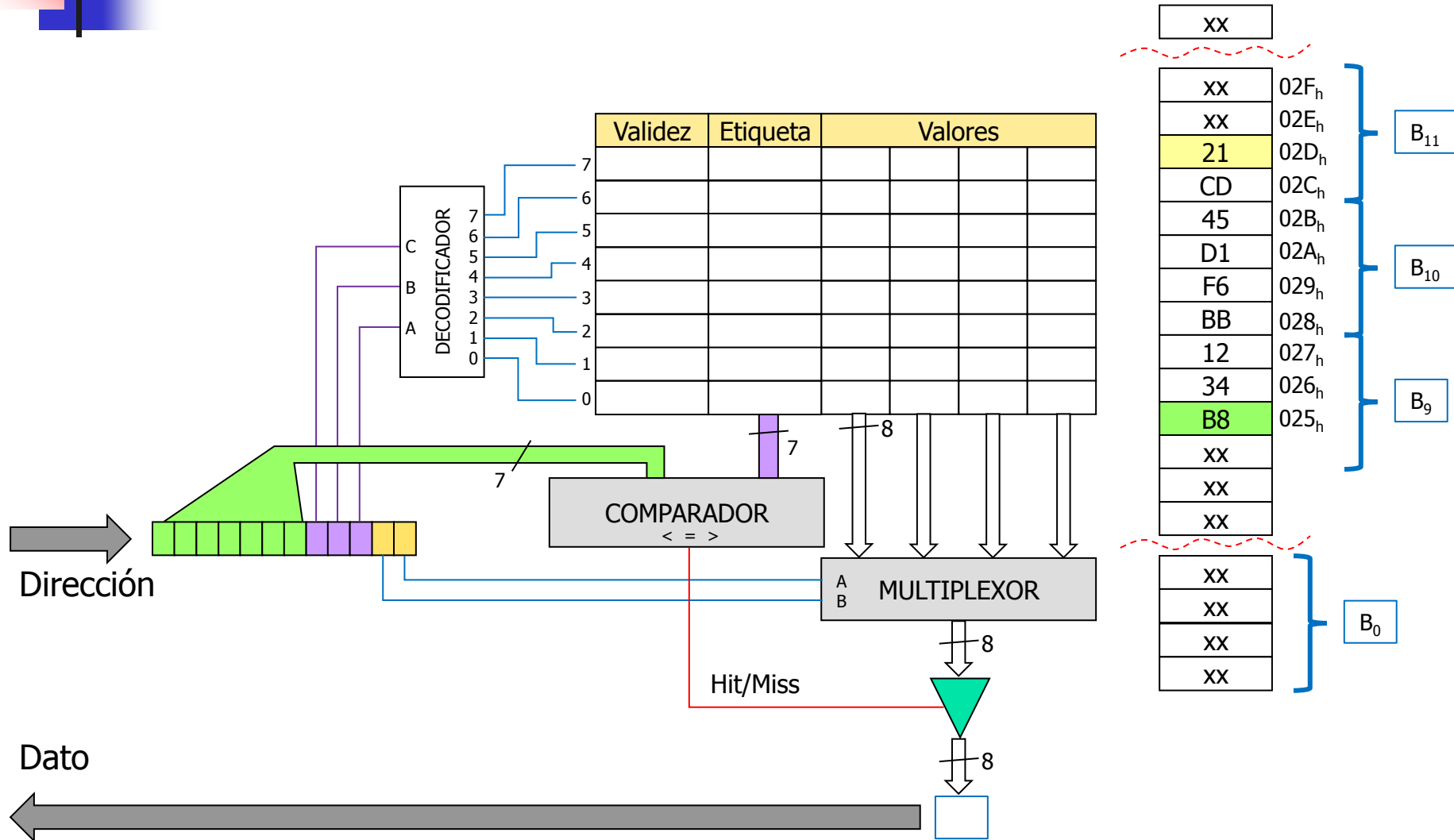
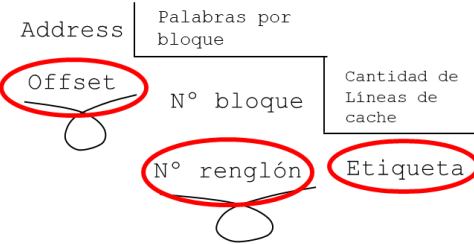
- No hace falta buscar

- La etiqueta resuelve el problema de mapeo de direcciones con igual número de renglón

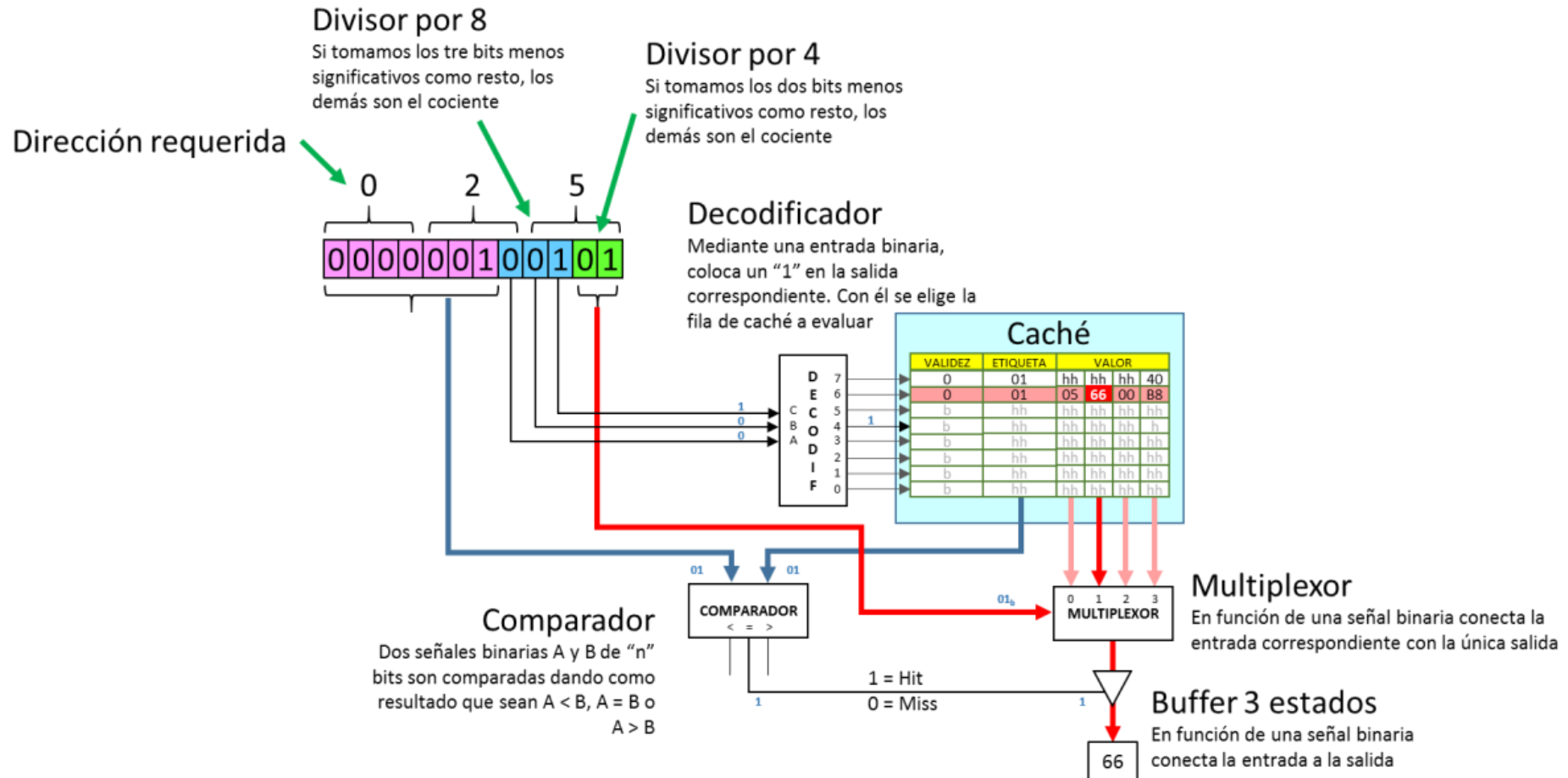
- Son más rápidas y sencillas



Organización mapeo directo

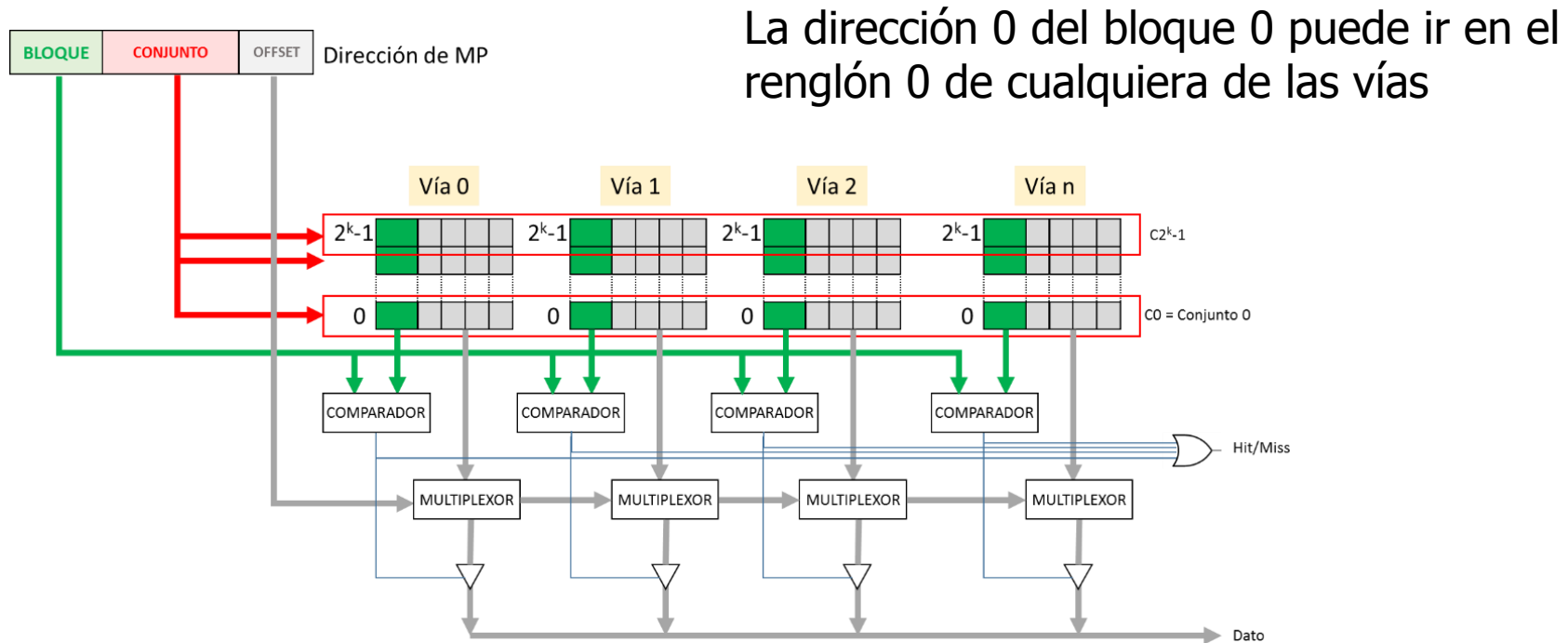


Organización mapeo directo



Organización asociativa por conjuntos

- Es una solución de compromiso entre ambas estrategias. Toma lo mejor de ambas para dar una mejor proporción de aciertos. Básicamente consiste en definir tamaño de bloques de MP de igual cantidad de palabras que cada paquete de direcciones o vía de memoria caché



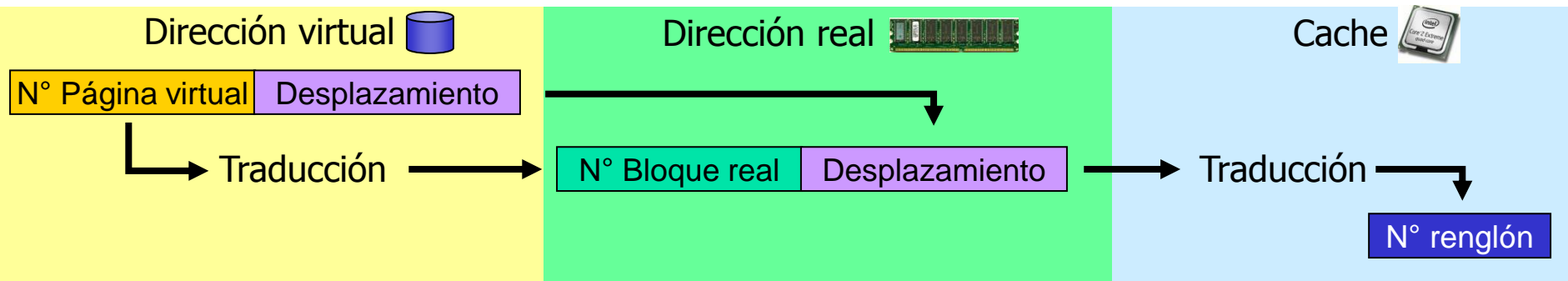


Memoria Cache

- Políticas de escritura
 - Determinan como se administrará las escritura de la memoria RAM común con los datos contenidos en la memoria Cache.
 - Escritura a Memoria (Write-Through Cache)
 - Cuando se modifica la Cache, se modifica el mismo dato en la memoria principal
 - No tiene problemas de consistencia
 - Beneficioso para procesos de lectura intensiva
 - Retrograbado (Write-Back Cache)
 - Se modifica solamente la Cache
 - Se modifica el dato en la memoria principal común cuando se necesita desocupar un bloque de la Cache que tiene datos modificados
 - Mas performante que el anterior
 - Beneficioso para procesos de escritura intensiva
 - Es necesario usar 1 bit más para indicar si el renglón de la Cache fue modificado o no – Dirty Bit
- ¿Qué bloque saco para que entre uno nuevo?
 - LRU (**L**east **R**ecently **U**sed) – RANDOM – FIFO (**F**irst-**I**n-**F**irst-**O**ut)

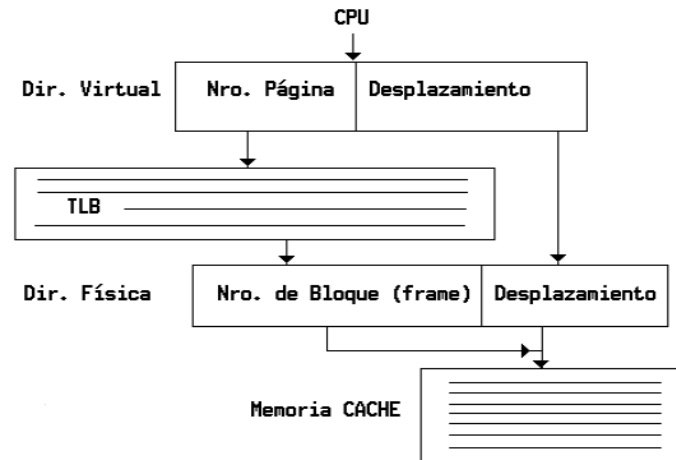
Memorias – Espacios

- Espacio Virtual
 - Definido en la memoria secundaria (discos externos, cintas, etc.)
 - La capacidad de direccionamiento en la memoria virtual supera las posibilidades definidas por la cantidad de bits del bus de direcciones (Address Bus)
 - Su capacidad está restringida a:
 - La capacidad de la memoria secundaria
 - La posibilidad de direccionamiento virtual de la CPU o del Sistema Operativo
- Espacio real
 - Definido en la memoria principal (MP)
 - La capacidad de direccionamiento está dada por la cantidad de bits del bus de direcciones
 - Su capacidad está restringida a la cantidad de memoria RAM física disponible
- Relación entre los diferentes espacios

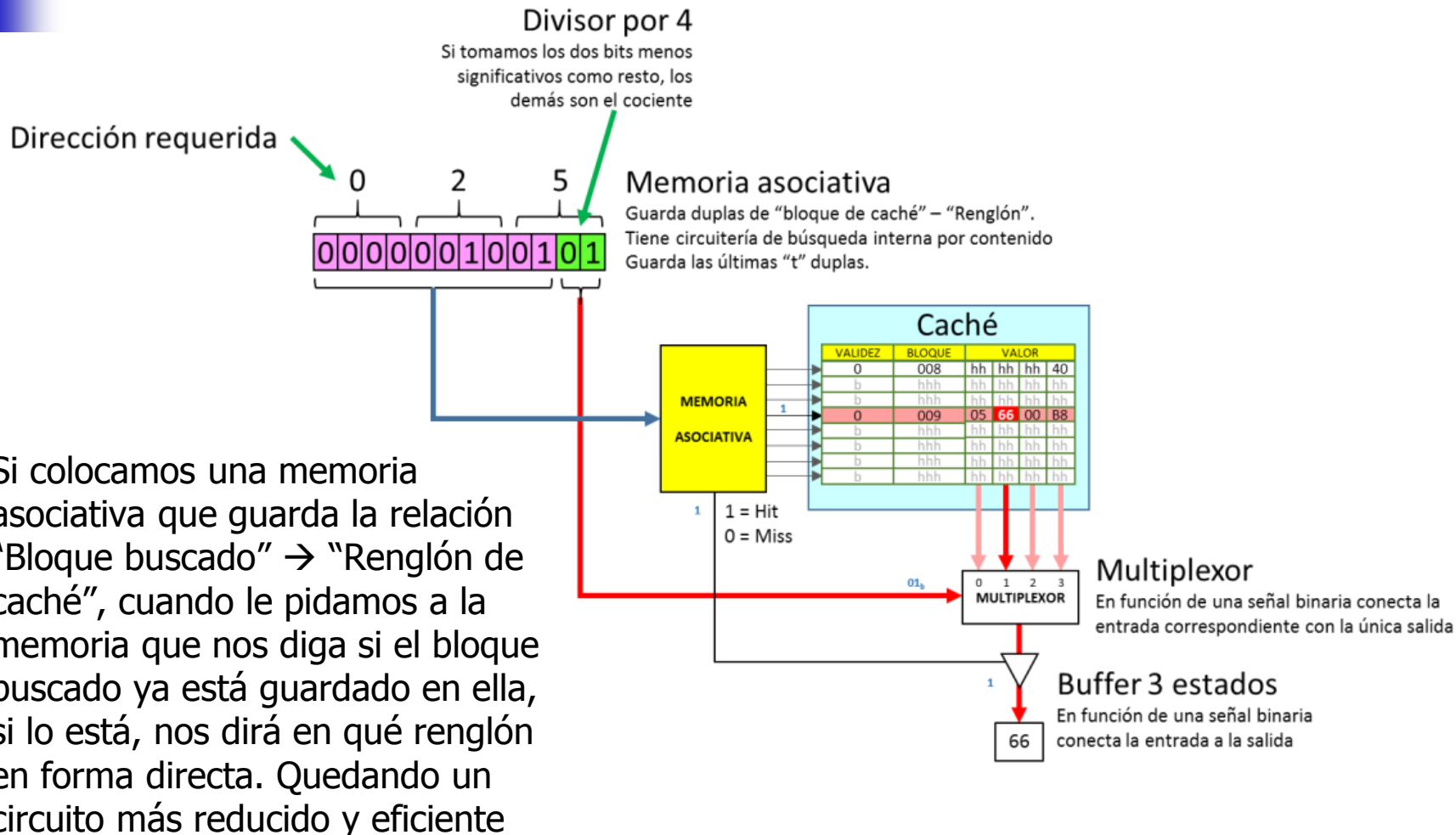


Memoria Cache

- Uso de la TLB (**T**able-**L**ookside-**B**uffer)
 - Las CPUs pueden direccionar espacios de memoria más allá de los n de bits (2^n) que posee en el bus de direcciones → Espacio virtual de memoria
 - La CPU traducirá esa dirección virtual de memoria (compuesta por un número de página + un desplazamiento) en una dirección real o física (compuesta por un número de bloque + un desplazamiento) para poder ubicarla en la memoria
 - La TLB es una memoria asociativa que almacena las últimas p traducciones
 - Si dirección virtual a traducir no está en la TLB, se resolverá a través de circuitería adicional

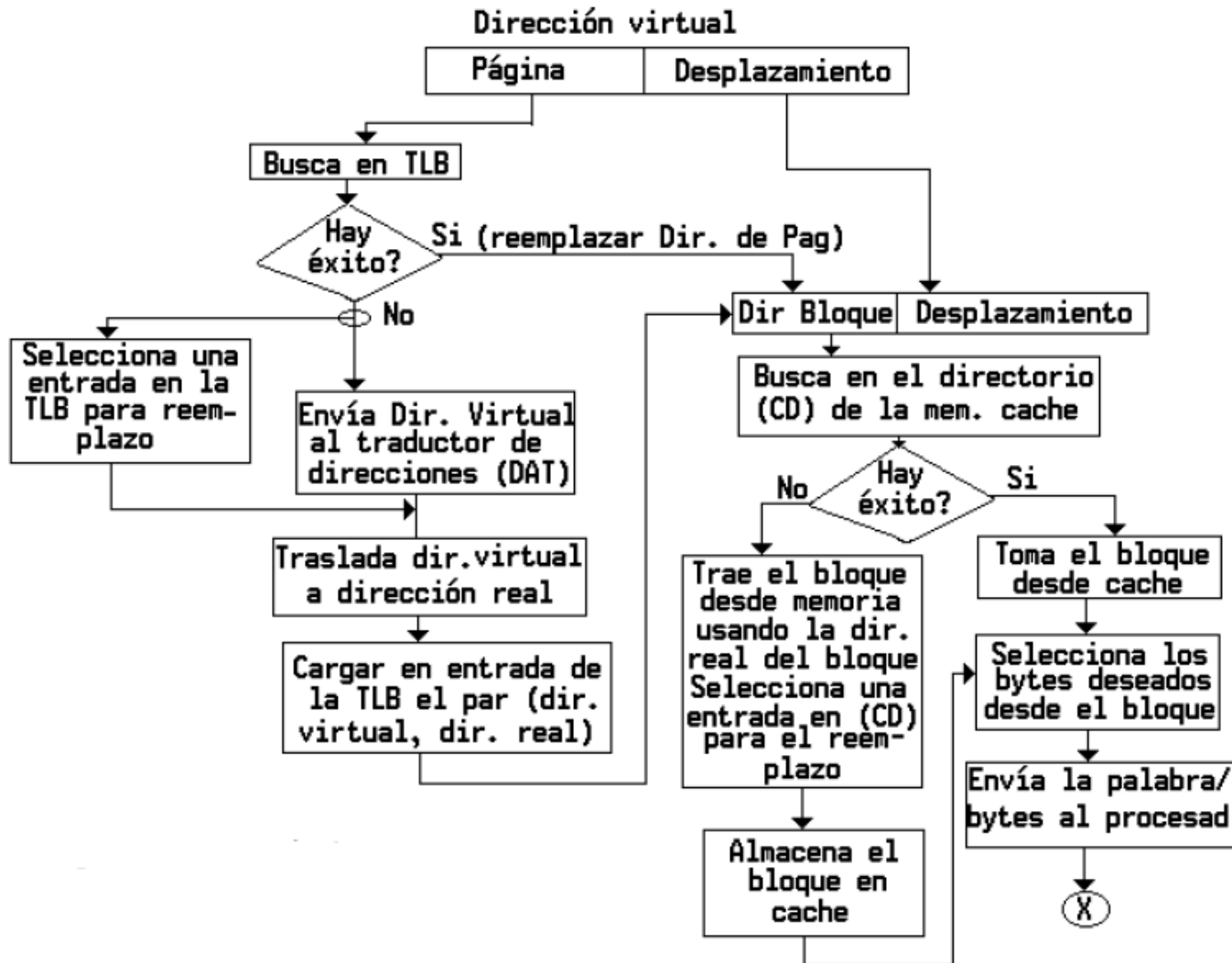


Organización asociativa – TLB



- Si colocamos una memoria asociativa que guarda la relación "Bloque buscado" → "Renglón de caché", cuando le pidamos a la memoria que nos diga si el bloque buscado ya está guardado en ella, si lo está, nos dirá en qué renglón en forma directa. Quedando un circuito más reducido y eficiente

Memorias – Espacios



Niveles de memoria caché

