



## Jerarquía de memorias (I)

- En las últimas décadas la velocidad de proceso de las UCP (medida en cantidad de instrucciones por segundo), se ha duplicado cada 18 meses sin variar su precio
- Las memorias han cuadruplicado su tamaño cada 36 meses, pero han tenido un progreso menos espectacular en velocidad, con incrementos de sólo alrededor del 10% anual
- Una computadora típica contiene distintos tipos de memoria, interactuando coordinadamente de modo tal que parece tener una única memoria grande y rápida.
- La forma en la que se organizan los distintos tipos de memorias se conoce como jerarquía de memorias.

Arquitectura de los Sistemas de Cómput

Jerarquía de Memorias

Registros
Cache
Principal
Discos Magnéticos

Cintas Magnéticas

Discos Magnéticas

Discos Magnéticas

Discos Opticos

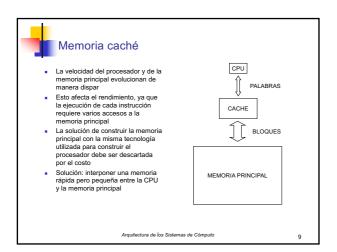
Jerarquía Tradicional

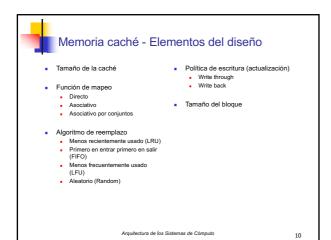
Arquitectura de los Sistemas de Cómputo

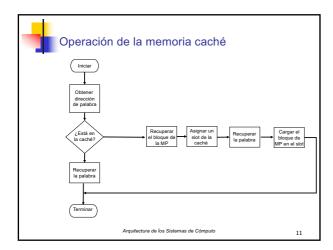


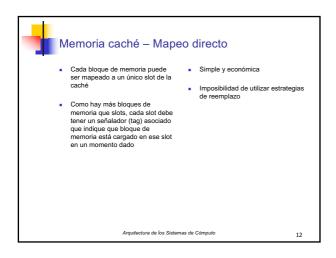


Categoría	Mecanismo de Borrado	Mecanismo de Escritura	Volatilidad	Otras	
Lectura /		de Lacilitara		Guuo	
	Eléctricamente		Volátil	Estáticas	
Escritura	(nivel de byte)			Dinámicas	
Lectura	Imposible	Máscaras	No volátil	Estáticas	İ
		Eléctricamente			
Princi- palmente de lectura	Radiación ultravioleta (nivel de chip)				
	Eléctricamente (nivel de bloque)				İ
	Eléctricamente (nivel de byte)				
	Princi- palmente de lectura	Radiación ultravioleta (nivel de chip) Electrura el electura Eléctricamente (nivel de bloque) Eléctricamente (nivel de byte)	Lectura Imposible  Radiación ultravioleta (nivel de chip) Princi- palmente le lectura Eléctricamente (invel de bloque) Eléctricamente Eléctricamente	Lectura Imposible  Radiación ultravioleta (nivel de tolio) Eléctricamente (nivel de bloque)  Eléctricamente (nivel de byte)	Lectura Imposible  Radiación ultravioleta (nivel de chio) Eléctricamente le lectura Eléctricamente (Eléctricamente Eléctricamente Eléctricamente (Eléctricamente Eléctricamente Eléctricam











#### Memoria caché – Mapeo asociativo

- Cada bloque de memoria puede ser mapeado a cualquier slot de la caché
- Cada slot de la caché debe mantener un señalador (tag) asociado que indique que bloque de memoria está cargado en ese slot de la caché en un momento dado.
- dado

  Dado que la búsqueda secuencial
  es ineficiente se debe implementar
  mediante una memoria asociativa
  (busca por contenido del señalador
  y no por dirección)
- Compleja y cara
- Permite la implementación de la estrategia de reemplazo más eficiente

Annuitantura de las Sistemas de Cémau

13

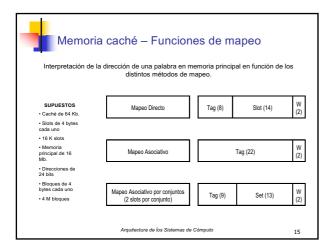


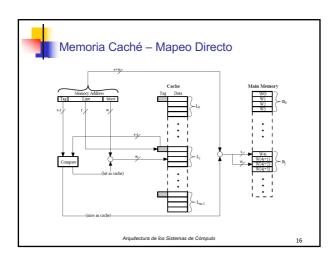
# Memoria caché – Mapeo asociativo por conjuntos

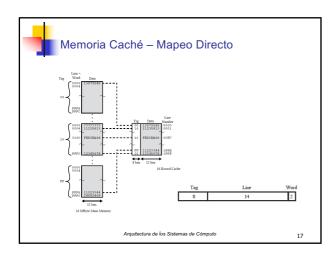
- La caché se divide en I conjuntos, cada uno de los cuales contiene J slots
- Un bloque de memoria es mapeado a un conjunto de slots de la caché y puede ser almacenado en cualquier slot del conjunto
- Cada slot de la caché debe mantener un señalador (tag) asociado que indique que bloque de memoria está cargado en ese slot de la caché en un momento
- Compleja y cara
- Permite la implementación de estrategias de reemplazo eficientes
- Normalmente el conjunto mantiene entre dos y cuatro slots

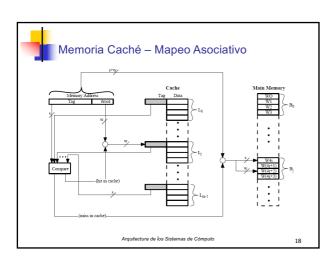
Arquitectura de los Sistemas de Cómputo

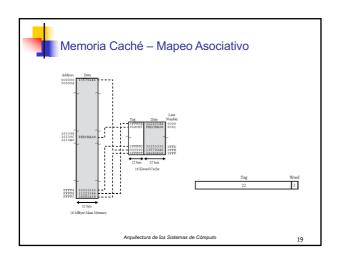
14

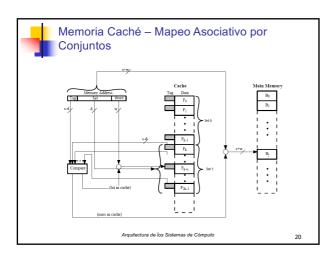


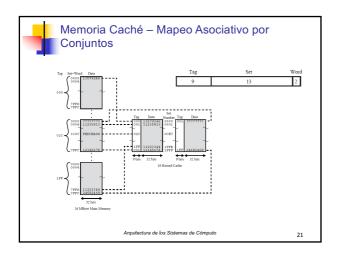














#### Memoria caché - Estrategias de reemplazo

- Mapeo Directo
  - Al haber un único slot en el que puede cargarse el bloque no hay posibilidad de implementar estrategias de reemplazo.
- Mapeos Asociativo y Asociativo por Conjuntos
  - LRU (Least-recently used): Reemplace el bloque que no ha sido referenciado por más tiempo
    - Se agrega a cada slot un bit de uso que se pone a uno cada vez que el slot es referenciado. Periódicamente todos se ponen a cero.
  - FIFO (First-in-first-out): Reemplace el bloque que hace más tiempo que está en la caché
    - Puede implementarse fácilmente con una técnica de buffer circular
  - LFU (Least-frequently used): Reemplace el bloque que ha sido usado menos veces
  - Se debe asignar un contador a cada slot que se pone a cero cuando se carga un nuevo bloque y se incrementa con cada referencia al bloque

    Random: Reemplace un bloque elegido al azar

22



#### Memoria caché - Políticas de escritura

- Se refiere a la política a seguir cuando un bloque que está en la caché es modificado
- Los problemas surgen en un equipo con DMA y en un contexto con múltiples procesadores
- Las políticas que normalmente se siguen son:
  - Write through: Cualquier actualización hecha a un bloque en la caché es inmediatamente realizada en el bloque de la memoria principal. Se asegura así que el contenido de la memoria principal siempre es válido.
  - Write back: Las actualizaciones son hechas exclusivamente en la caché y sólo son transferidas a la memoria principal cuando el bloque es reemplazado en la caché. Por momentos el contenido de la memoria principal puede ser inválido.

Arquitectura de los Sistemas de Cómputo



## Memoria caché - Coherencia

- Es un problema que surge en una organización en bus cuando los dispositivos conectados al bus tienen cada uno su propia caché.
- Si una palabra en una caché es modificada, se invalida no sólo el contenido de la memoria principal sino también el de otras cachés. Las políticas que normalmente se siguen son:
- - Bus Watching with Write Through: Cada caché monitorea el bus para detectar referencias de escritura a palabras de memoria que están en ellas e invalidarlas cuando así ocurre
  - Hardware transparency: Se agrega hardware para asegurar que las actualizaciones a la memoria principal vía una caché se reflejan en todas las cachés
  - Non-cacheable Memory: Sólo una porción de la memoria principal es compartida por más de un procesador y esa porción de memoria no puede llevarse a la caché.

Arquitectura de los Sistemas de Cómputo



#### Memoria caché – Tamaño del bloque

- Bloques más grandes reducen el número de bloques que entran en la caché, lo que puede obligar a su reemplazo poco después de haber sido cargados.
- Cuando el tamaño del bloque crece cada palabra adicional está más lejos de la palabra requerida que provocó la carga en la caché y, en consecuencia, disminuye la probabilidad de que sea utilizada en el futuro cercano.
- No hay un tamaño óptimo, ya que éste depende de las características de localidad de cada proceso particular
- Normalmente se usa un tamaño de entre 4 y 8 unidades direccionables (palabras o bytes)

25



#### Discos Magnéticos

- Son platos construidos de metal o plástico recubiertos con un material magnetizable.
- Los datos son grabados o leídos mediante una *cabeza* de lectura escritura
- Durante las operaciones de lectura la cabeza está estacionaria y el plato gira debajo de ella

- Conceptos
  Pista (track)
  Cilindro (cylinder)
  Inter Track Gap (ITG)
  Sectores o bloques

  - Sectores o bloques
     Longitud fija
     Longitud variable
     Inter Block Gap (IBG)
     ID Field
     SYNCH byte

  - Densidad (bits per inch)
    Tiempo de acceso
    Demora de rotación
    Velocidad de transferencia
  - Factor de intercalado

Arquitectura de los Sistemas de Cómputo



# Discos Magnéticos - Características

- Movilidad de las cabezas de lectura Platos

  - FijasMóviles
- Portabilidad
  - Removibles
     No removibles (fijos)
- Mecánica de la cabezas de lectura
  - / escritura

    Contacto (floppy)
  - Separación fija Separación aerodinámica
- UnoVarios (disk pack)
- Lados
  - Simple (uno solo)Doble (ambos)

Arquitectura de los Sistemas de Cómputo

