


# Arquitectura de Computadores

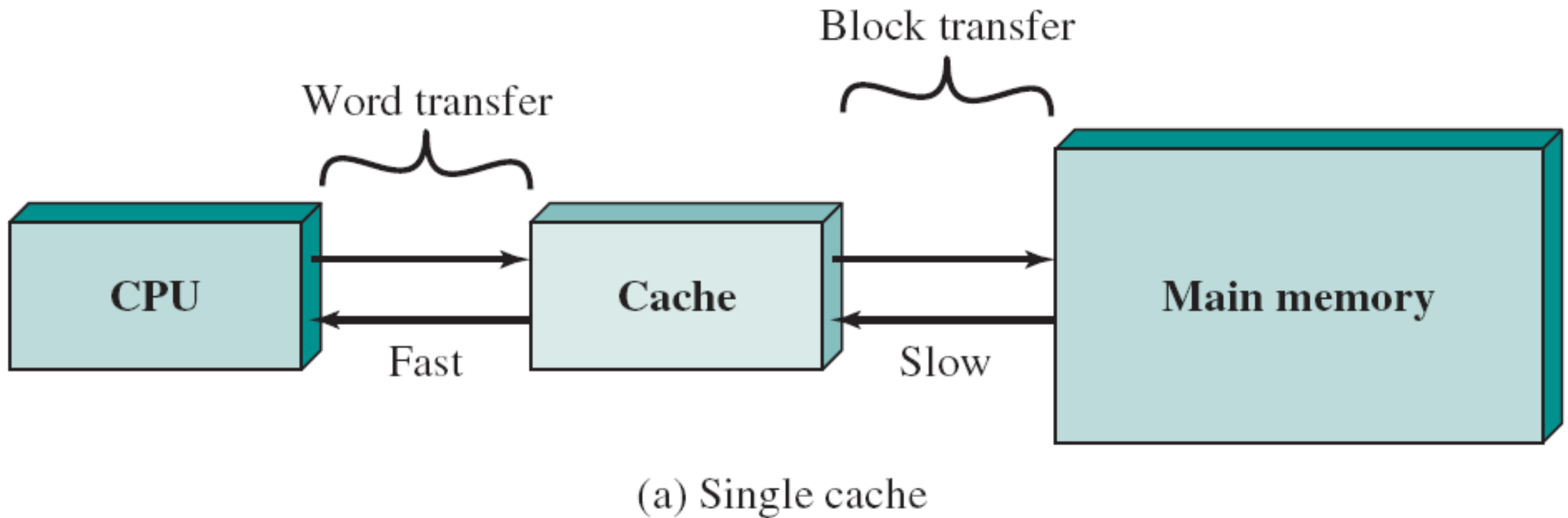
The background of the slide is a collage of digital and technological motifs. On the left, a blue silhouette of a human head is filled with a circuit board pattern, with binary code (0s and 1s) floating around it. The right side features a dark blue background with glowing white circuit traces and bright light flares. In the top right corner, there is a faint, stylized neural network diagram.

2021 A

Andrés Larco

# MEMORIA CACHE

## 6.2 Principios de funcionamiento de la cache



# Introducción

- El funcionamiento de un computador depende de la interacción entre la CPU y el sistema de memoria.
- Un sistema de memoria perfecto es aquél que puede suministrar ***inmediatamente*** cualquier dato que la CPU solicite.
- No obstante, esta memoria ideal no es implementable en la práctica. No existe ningún dispositivo que combine adecuadamente la capacidad, la velocidad y el coste (tamaño).



# Jerarquía de memoria

Se pretende que el procesador **perciba** una memoria de la velocidad del nivel más próximo con la capacidad del nivel más lejano.

La *cache* se introduce para completar esta percepción.

La cache es implementada en 2 niveles.

# Jerarquía de memoria

---

Cada nivel de memoria contiene una copia de un subconjunto de los datos del siguiente nivel.

---

El último nivel contiene la totalidad de datos almacenados en memoria

---

El número de niveles no está restringido.

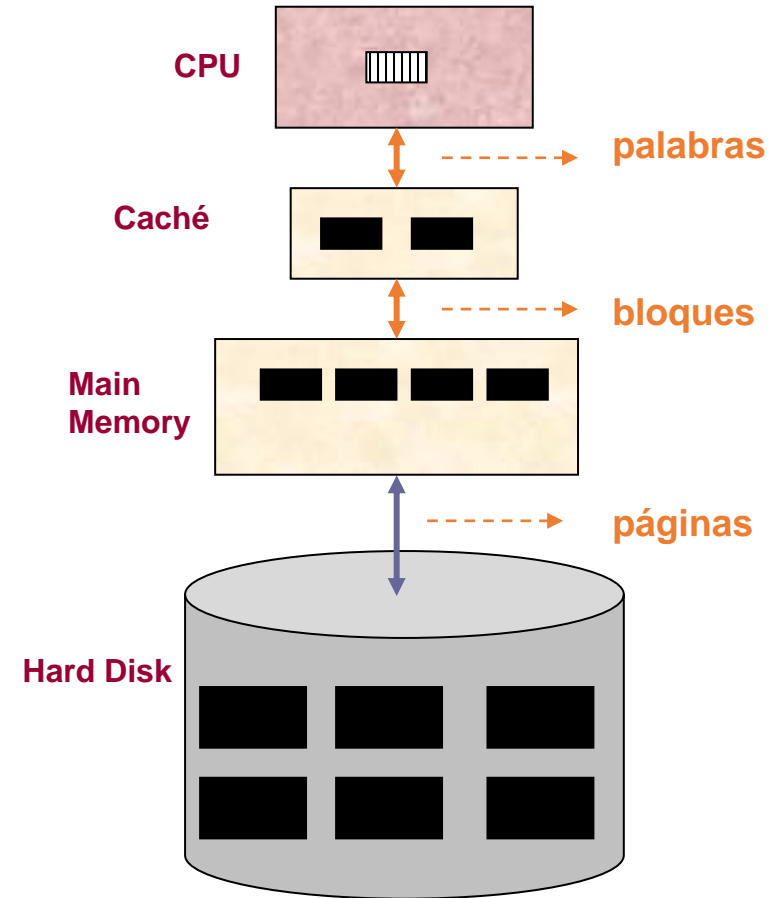
---

El procesador opera directamente con el nivel 0 de memoria.

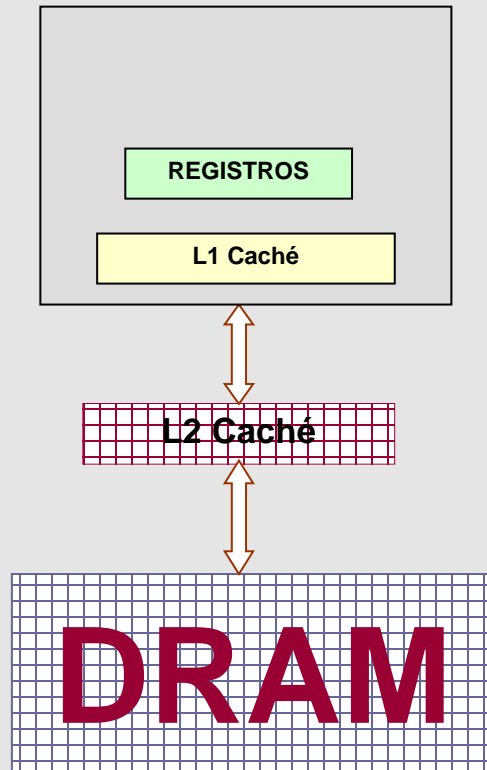
---

Se tiene que diseñar el movimiento de datos entre niveles.

---



# ¿Qué es la memoria Caché?



- Esta es una memoria pequeña y rápida que se interpone entre la CPU y la memoria principal para que el conjunto opere a mayor velocidad.
- Se basa en un principio de uso común en la vida diaria:

*“poner más a la mano las cosas de uso más frecuente”*

# La memoria Caché

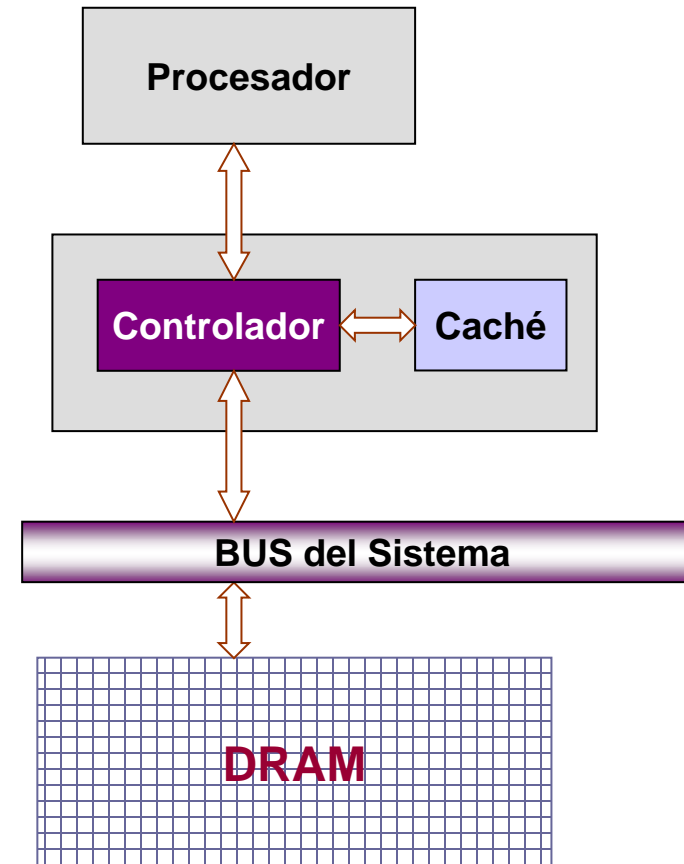
Se basa en dos suposiciones que generalmente resultan ciertas:

- **Los computadores tienden a utilizar las mismas instrucciones y los mismos datos repetidamente.**
- **La información necesitada se encuentra almacenada de forma adyacente, o cuando menos muy cercana.**

Así, la caché contiene *una copia de una parte* de los datos e instrucciones que están en memoria principal.

# La memoria Caché

- Esta copia debe estar disponible justo cuando el procesador la necesita permitiéndole acceder a esos ítems sin recurrir a estados de espera.
- Requiere de **hardware adicional** que asegure que este pequeño banco de memoria caché cumpla con su función:
  - Controlador (chipset)
  - buses



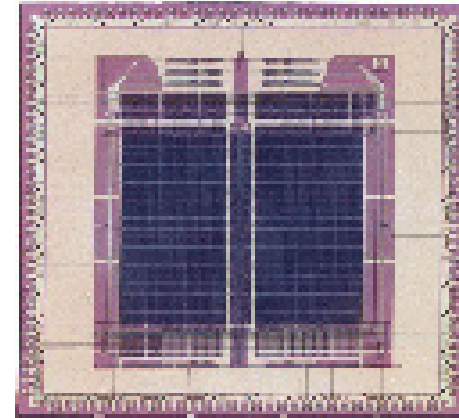


# Tecnología: SRAM

- Almacena la información en un biestable.
- Una celda (un bit) se compone de seis transistores, por lo que presenta menor capacidad de almacenamiento por chip.
- Tres transistores consumen energía máxima en forma permanente y los otros tres consumen mínima energía, por lo que presenta un mayor consumo.
- La lectura es directa y no destructiva, así su tiempo de acceso es muy bajo

# SRAM Características

- Memoria de acceso aleatorio.
- Memoria estática, es decir mantiene todos sus datos sin necesidad de refresco externo.
- Volátil, es decir mantiene los datos si existe el suministro de energía hacia el circuito.



**SRAM**

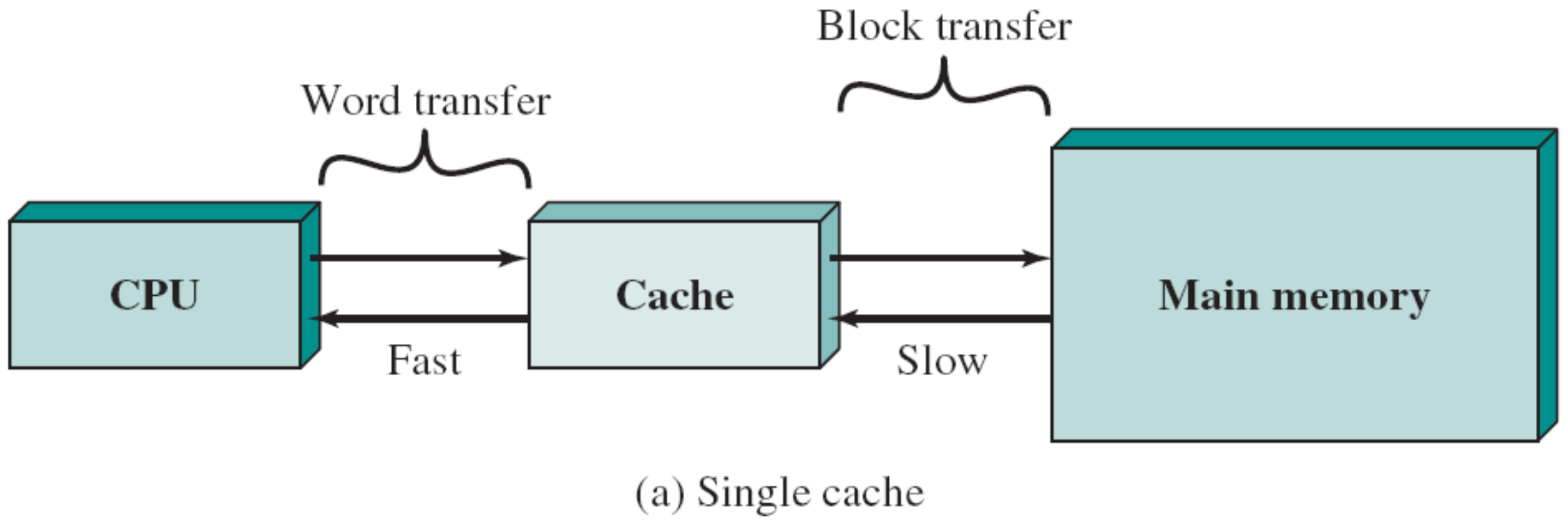
# Síntesis

- Pequeño dispositivo que forma parte del sistema de memoria, ocupa el segundo lugar en la jerarquía.
- Contiene un subconjunto de los datos e instrucciones que están almacenados en la memoria principal.
- La idea es que el micro acceda lo más rápidamente a la información que requiera (incluso sin recurrir a estados de espera).

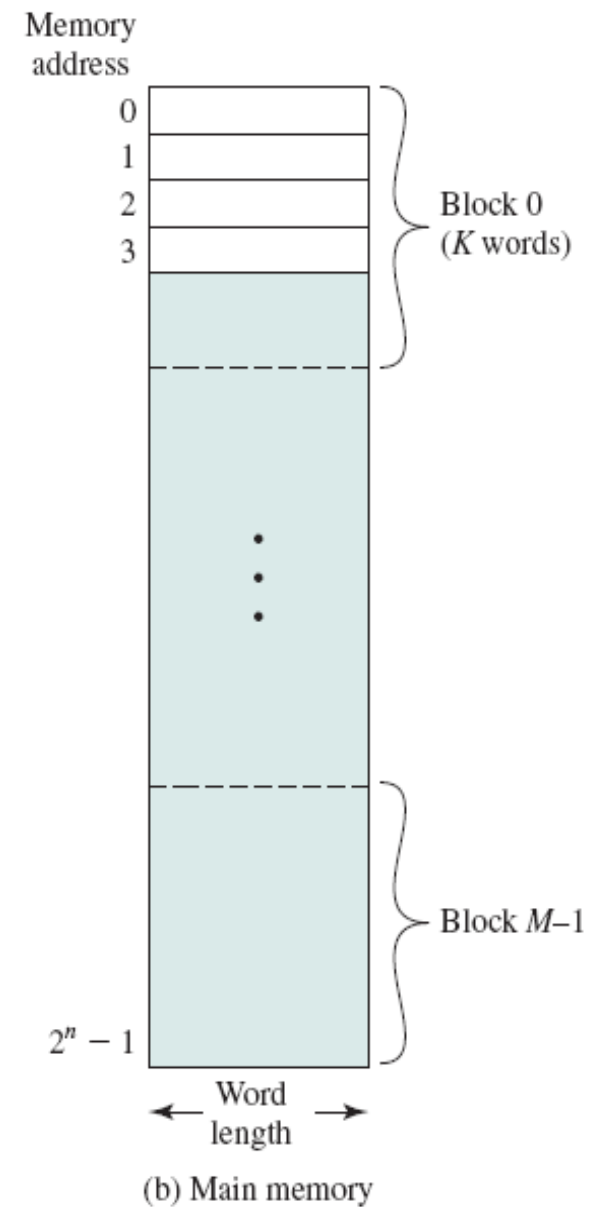
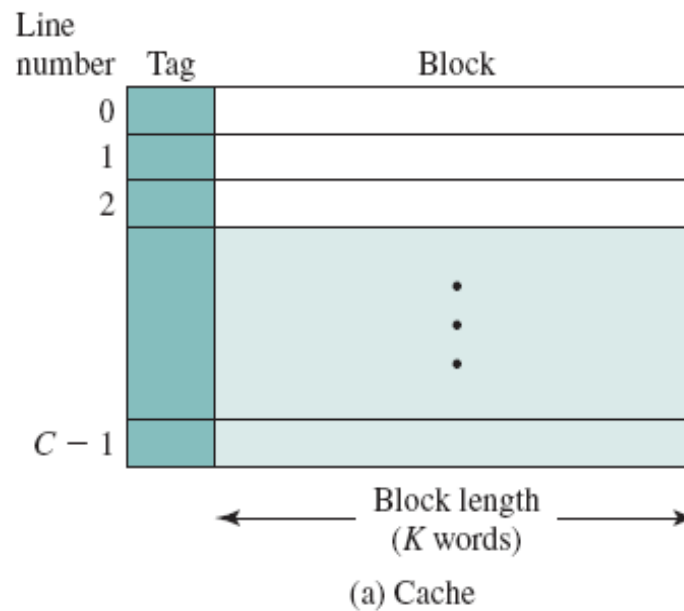
# Síntesis

- Físicamente es un chip SRAM, con las siguientes características:
  - Acceso aleatorio
  - Acceso por dirección / contenido
  - Lectura / Escritura
  - Volátil
  - Estática
  - Es una memoria NDRO
  - Tiempo de acceso bajo comparado con la DRAM
  - Se estructura físicamente en base de circuitos biestables (6 transistores, lo que les hace más caras y más grandes.
  - Tres ubicaciones diferentes: L1 y L2

# Cache y memoria principal

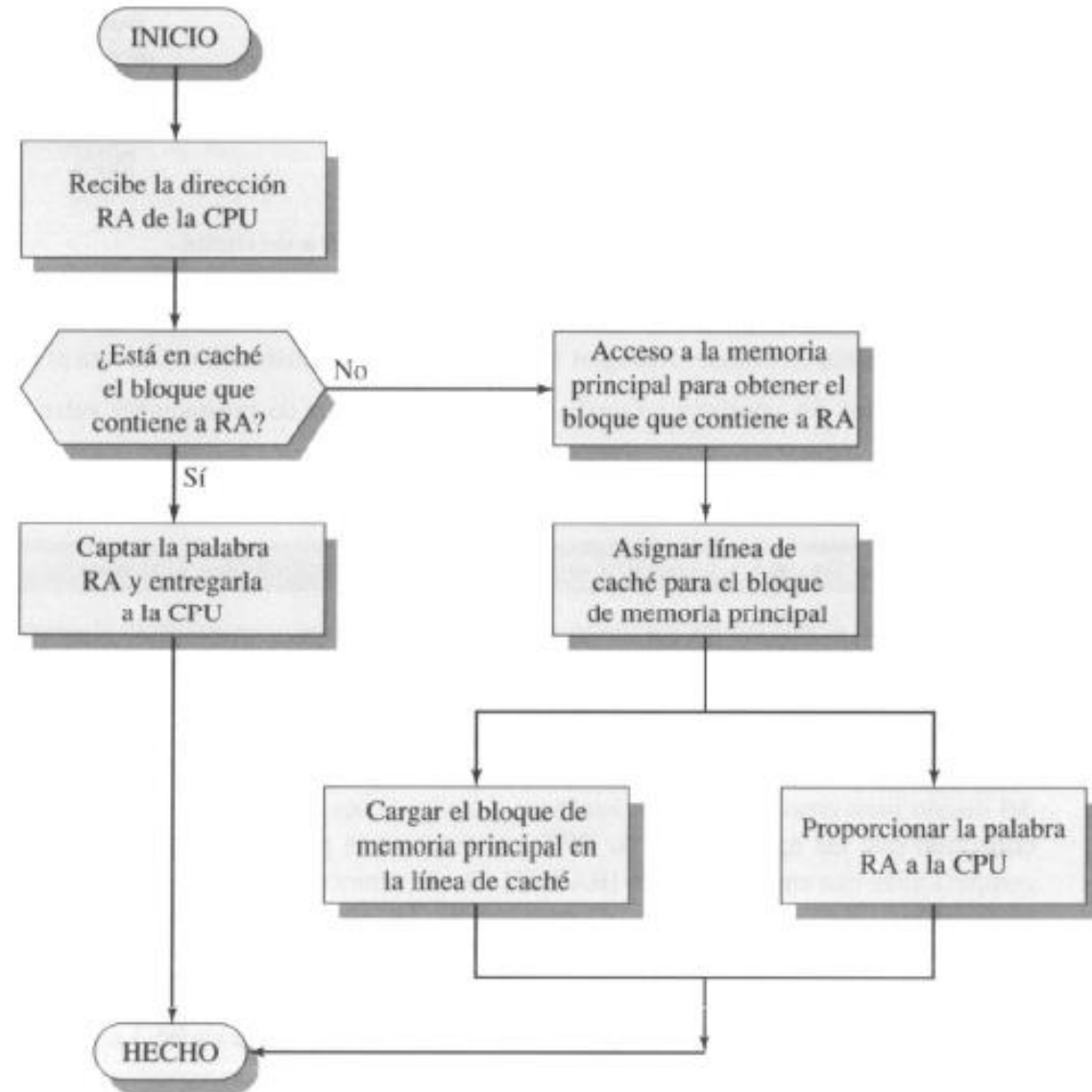


# Estructura de la Cache y la memoria principal

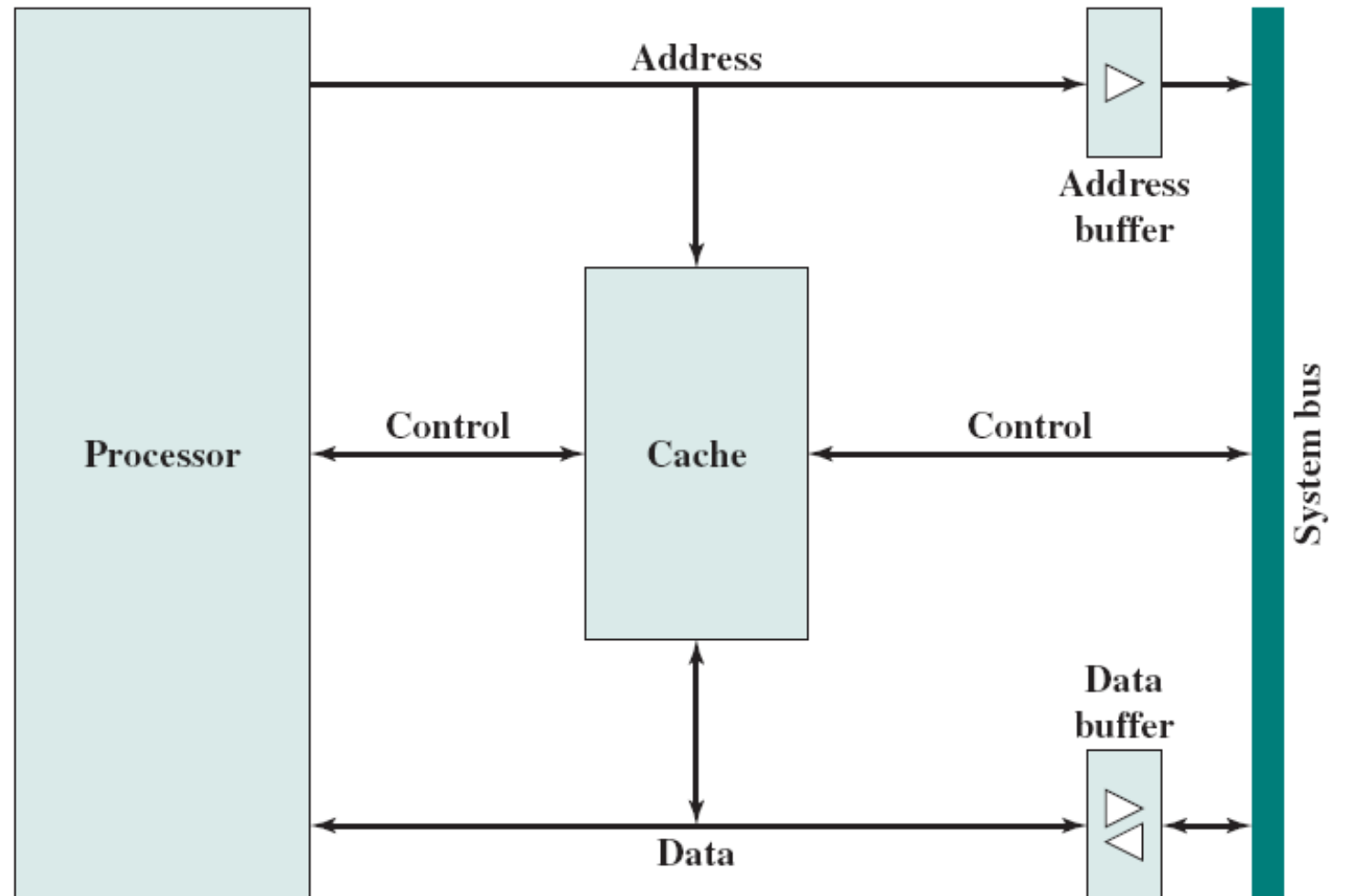




# Operación de lectura de la cache



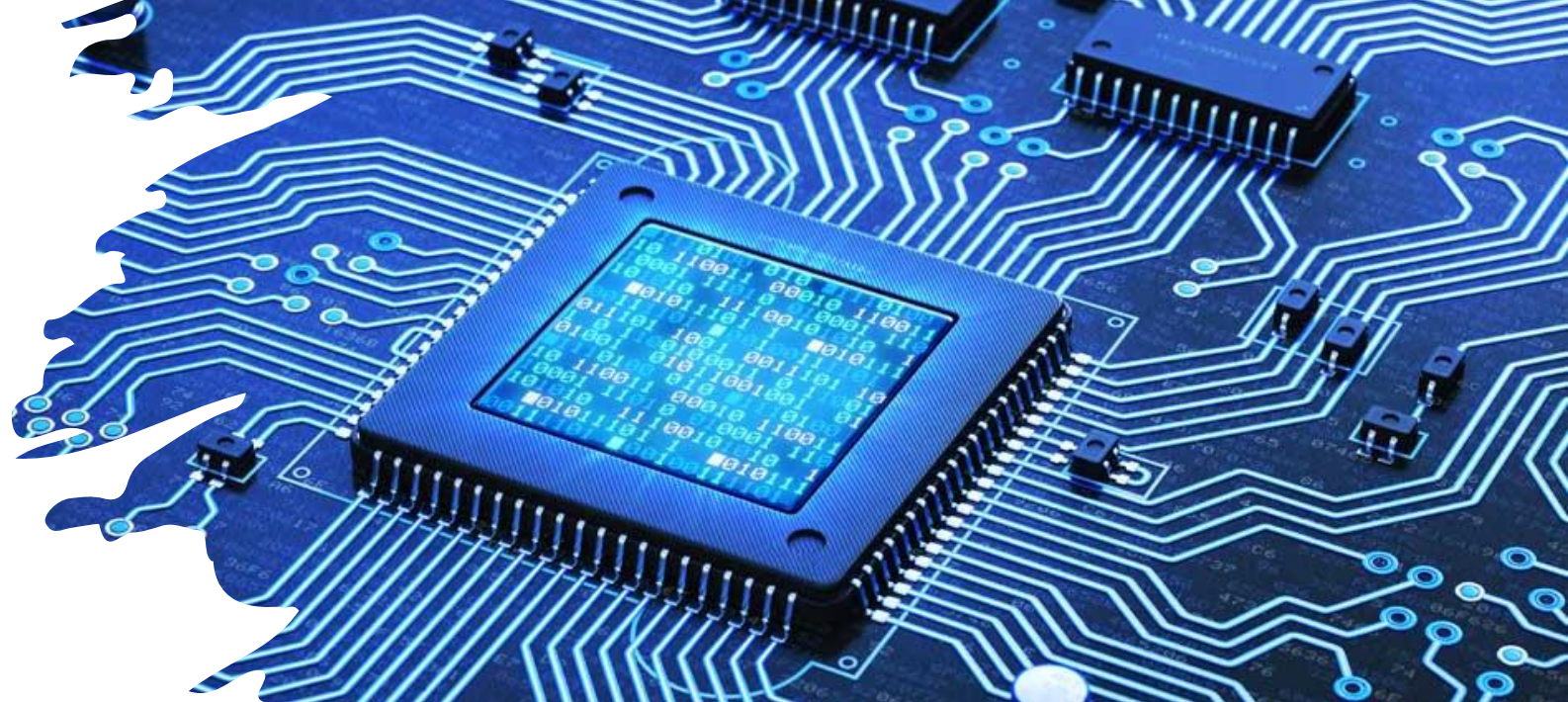
# Organización típica de la cache



# Referencias:

---

William Stallings, Computer Organization and Architecture:  
Designing for Performance, 9th Edition, Prentice Hall, 2010, ISBN-  
13: 978-0-13-607373-4



Gracias !!!

