

# 计算机组成原理与系统结构

## 第五章 存储体系

[http://www.icourses.cn/coursestatic/course\\_2859.html](http://www.icourses.cn/coursestatic/course_2859.html)





## 5.5 高速缓冲存储器 Cache

主存与 Cache 的地址映射方

式

Cache 的  
基本原理

替换算法

Cach  
e

写策略

Cache 的多层次设  
计





# 一、Cache 的基本原理



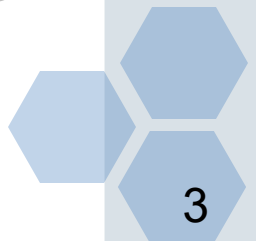
Cache 的  
特点



Cache 的  
工作原  
理



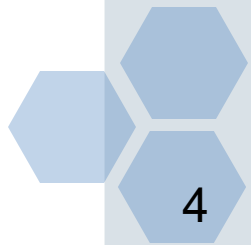
Cache  
的命中  
率





# 1、Cache 的特点

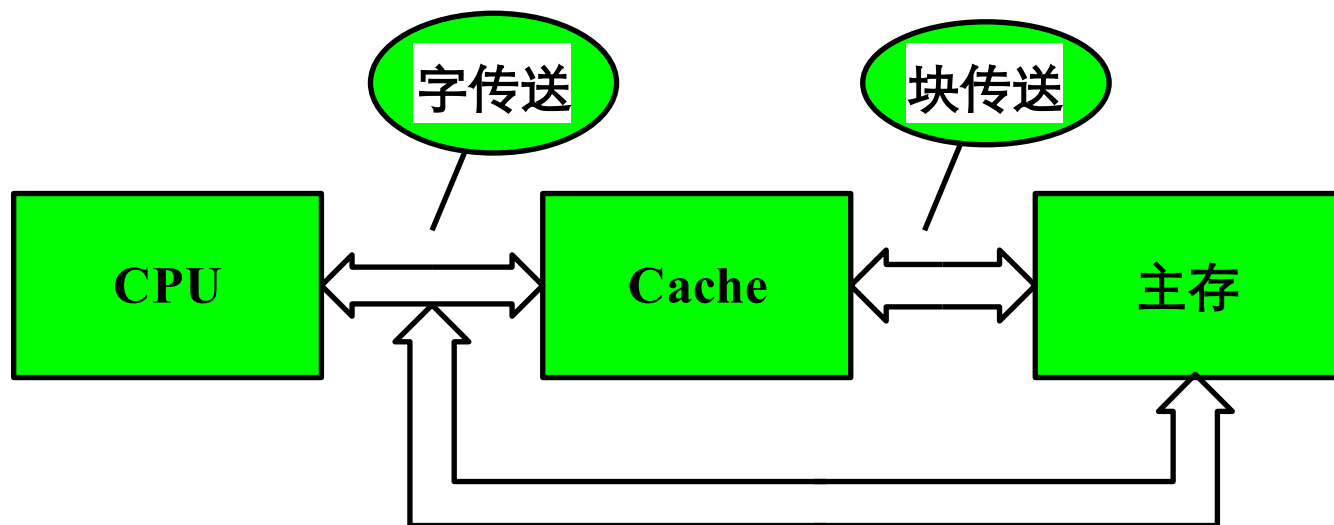
- Cache 是指位于 CPU 和主存之间的一个高速小容量的存储器，一般由 **SRAM** 构成。
- Cache 功能：用于弥补 CPU 和主存之间的**速度差异**，提高 CPU 访问主存的平均速度。
- 设置 Cache 的理论基础，是**程序访问的局部性原理**。
- Cache 的内容是主存部分内容的**副本**，Cache 的功能均由**硬件**实现，对程序员是**透明的**。





## 2、Cache 的工作原理

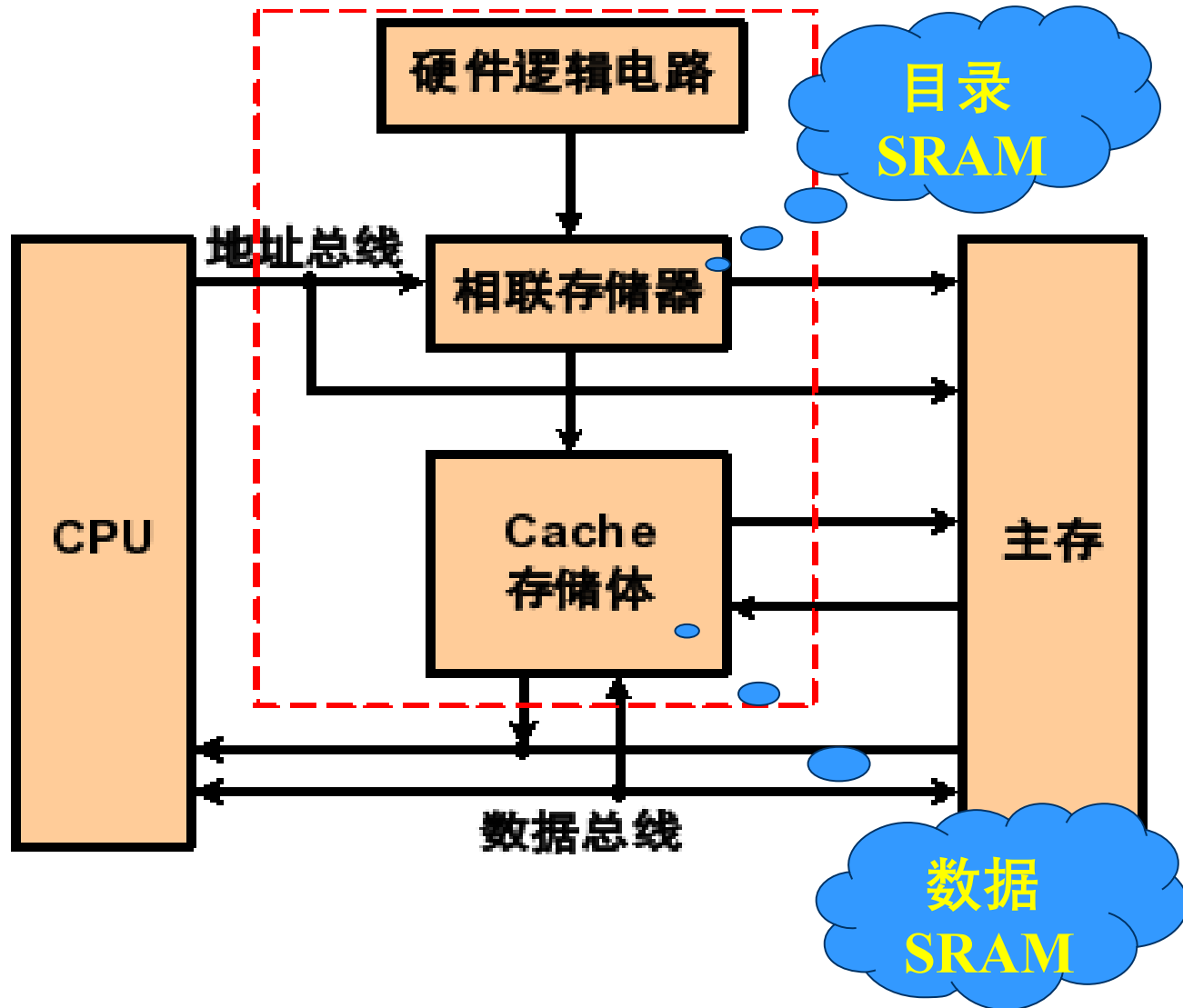
■ Cache 的速度比主存快 5 — 10 倍。



Cache、主存与 CPU 的关系



# Cache 的原理图





# Cache 的读写操作

1. CPU 在读写存储器时，Cache 控制逻辑首先要依据地址来判断这个字是否在 Cache 中，若在 Cache 中，则称为“命中”；若不在，则称为“不命中”。
2. 针对命中 / 不命中、读 / 写操作，Cache 的处理是不同的：
  - 读命中：立即从 Cache 读出送给 CPU ；
  - 读不命中：通常有两种解决方法：
    - ① 将主存中该字所在的数据块复制到 Cache 中，然后再把这个字传送给 CPU ；



# Cache 的读写操作

②把此字从主存读出送到 CPU ，同时，把包含这个字的数据块从主存中读出送到 Cache 中。

■ 写不命中：直接将该字写入主存中，且不再调入 Cache ；

■ 写命中：通常也有两种方法进行处理：

①写贯穿方法：同时对 Cache 和主存进行写操作；

②写回：只写 Cache ，仅当此 Cache 块被替换时，才将该字写入主存

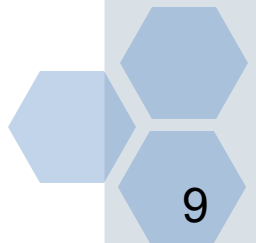






### 3、Cache 的命中率

1. **命中率**指 CPU 访问主存数据时，命中 Cache 的次数，占全部访问次数的比率；失效率就指不命中 Cache 的次数，占全部访问次数的比率。命中率  $h$  取决于程序的行为、Cache 的容量、组织方式、块大小。





### 3、Cache 的命中率

2. 在一个程序执行期间，设  $N_c$  表示 Cache 完成存取的总次数， $N_m$  表示主存完成存取的总次数，则命中率：

$$h = \frac{N_c}{N_c + N_m} \times 100\%$$

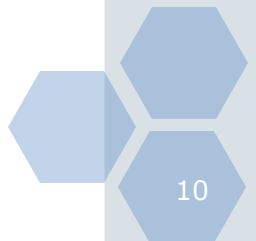
若  $t_c$  表示 Cache 的访问时间， $t_m$  表示主存的访问时间，则 Cache/ 主存系统的平均访问时间  $t_a$  为：

$$t_a = ht_c + (1-h) \times (t_m + t_c)$$

$$t_a = ht_c + (1-h) t_m$$

Cache/ 主存系统的访问效率  
 $e$ ：

$$e = \frac{t_c}{t_a}$$





## 二、主存与 Cache 的地址映射方

❖ **讨论的问题：** 如何根据主存地址，判断 Cache 有无命中并变换为 Cache 的地址，以便执行读写。  
有三种地址映射方式：

- 1 直接映射
- 2 全相联映射
- 3 组相联映射

❖ **讨论前提：** Cache 的数据块称为行，主存的数据块称为块，行与块是等长的；主存容量为  $2^m$  块，Cache 容量为  $2^c$  行，每个字块中含  $2^b$  字。





# 1、直接映射

❖ **特点：** 是一种多对一的映射关系：主存的第  $i$  块一定映射到 Cache 的第  $j$  行，且：

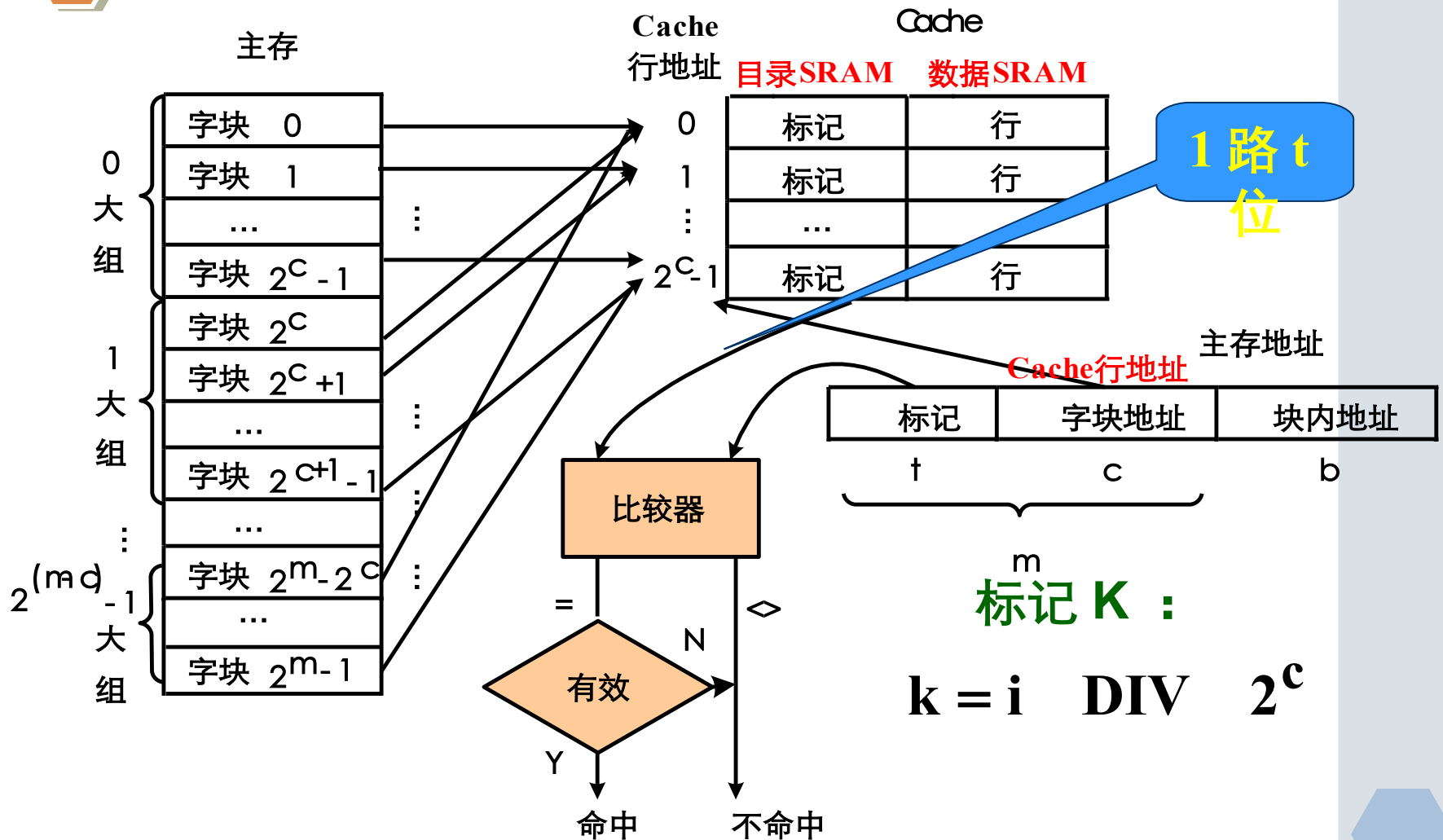
$$j = i \bmod 2^c$$

① 优点：映射方式简单，易实现。

② 缺点：机制不灵活，Cache 命中率低。



# 1、直接映射





## 直接映象法举例

❖ 例如：主存 4GB（按字节编址），Cache 512KB，块大小 16B；读主存 234589ABH 的过程（假设 Cache 空白）。

主存划分格式

主存地址：\_\_ 32 \_\_ 位

块大小  $16\text{B}=2^4\text{B}$ ，按字节编址，所以块内地址：\_\_ 4 \_\_ 位

Cache 分为：\_\_  $512\text{KB} \div 16\text{B} = 2^{15}$  \_\_ 行（块）

标记	Cache 行（块）地址	块内地址
13 位	15 位	4 位
0010 0011	101 1000 1001	1011

