

# 存储体系

一个二进制位是存储器最小单位；字节是数据存储的基本单位  
单元地址是内存单元的唯一标志

存储字：存储芯片中的一个读写单位

编址单位：一个存储单元的位数，现在都按字节编址

传输单位：

- 一次从主存读出或写入的位数，可以**不等于**存储字的长度，也可以**不等于**编址单位
- 对外而言，数据通常按**块传输**，传输单位为块

## 存储器的分类

### 按存储介质分类

- 半导体器件（用做主存）
  - RAM
  - ROM
- 磁性材料（用做辅存）
  - 磁带
  - 磁盘
- 光介质（光盘存储器）

## 随机读写存储器RAM

### 静态存储器（SRAM）

#### 地址译码方式

- 线性译码方式：n位地址线，经过一维译码后，有 $2^n$ 根选择线
- 双向译码方式：n位地址分成行列地址有 $2^{(n/2)+1}$ 根选择线  
使得选择线大大减少

#### 特点

- 使用双稳态触发器表示0和1代码
- 电源不掉电的情况下，信息保持稳定
- 存取速度快，集成度低（容量小），价格高
- 抗干扰性好

- 常用作高速缓冲器Cache

## 动态存储器 (DRAM)

- 刷新操作：按**行**来执行内部读操作，由刷新计时器来生成行地址，读出方式刷新，**刷新一行所需的时间就是一个存储周期**
- 刷新周期：从对整个存储器刷新结束时起，到对整个DRAM全部刷新一遍为止的时间间隔
- 刷新信号时间：相邻两个行位元之间刷新的时间间隔
- 单元刷新时间间隔：DRAM允许**最大信息保持时间**，一般为64ms
- 刷新行数：单个芯片的单个矩阵行数

## 特点

- 采用双向译码（容量大）
- 芯片地址引脚：
  - 约为地址码位数的一半
  - 行、列地址分时输入
- 控制引脚：行选通引脚和列选通引脚
- 含刷新和再生电路
- 一定先读行
- 半导体器件中分布电容上有无电荷来表示1和0
- 电源不掉电的情况下，信息也会丢失，因此需要不断地刷新
- 存取速度慢，集成度高（容量大），价格低
- 和SRAM相比，抗干扰性差
- 常用作内存条

每出现新一代存储芯片，容量至少提高4倍数

## 刷新方式

### 集中式刷新

在内部刷新时间内不允许访存，这段时间称为**死时间**

### 分散式刷新

在任何一个**存储周期**内，分为**访存**和**刷新**两个子周期

- 访存时间内，供CPU和其他主设备访问

- 在刷新时间内，对DRAM的某一行进行刷新  
**系统存储周期为存储器周期的两倍，即 $500\text{ns} \times 2$**   
**刷新周期缩短**

### 异步式刷新

折中的办法，在64ms内分散地把各行刷新一遍

避免分散式刷新中不必要的多次刷新，提高整机速度；同时又解决了集中式刷新中“死区”时间过长的问題

## 按存储方式分类

- 顺序存取存储器：存取的时间和存储单元的物理位置有关（磁带）
- 直接存取存储器：利用一个共存读写机制，直接定位到要读写的数据块，在读写某个数据块时按顺序进行（磁盘）
- 相联存储器：按内容访问

## 按存储器的读写功能分类

- 只读寄存器（ROM）
- 读写寄存器（RAM）

以上两个寄存器一般隐含随机存取

## 只读寄存器ROM

### 特点

- 信息只读不能写
- 非破坏性读出，无需再生
- 可以进行随机存取

## 按信息的可保存性分类

- 永久记忆的存储器：在断电后还能保存信息（辅存，ROM）
- 非永久记忆的存储器：断电后信息丢失（主存的RAM）

## 按在计算机系统中的作用分类

### 寄存器（Register）

封装在CPU内，用于存放**正在执行的指令和使用**的数据

## 高速缓存 (Cache)

位于CPU内部或附近，用来存放**要执行**的局部程序段和数据  
用SRAM实现，速度快，容量小

软件人员感觉不到Cache的存在，称为Cache的**透明性**

### Cache命中

Cache的地址存放在一个相联存储器中，当CPU提供字地址时按内容进行查找，如果地址匹配成功则称为Cache命中。反之称为Cache不命中或Cache缺失。

### 命中率

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

$N_e$ 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

$h$ 定义位命中率

### 平均访问时间

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

$t_c$ 表示命中时Cache的访问时间， $t_m$ 表示Cache缺失时主存的访问时间

### 访问效率

$$e = \frac{t_c}{t_a} = \frac{t_c}{ht_c + (1 - h)t_m}$$

### Cache缺失的处理方法

- 将主存中该字**所在的行**复制到**Cache**中，再由**Cache把这个字传给CPU**
- 启动常规的主存读周期，把该字**从主存读出送到CPU**，同时把包含这个字的行**从主存中读出后写入Cache**。

### 主存与Cache的地址映射方式

Cache的数据块大小称为**行**，主存数据块大小称为**块**，**行与块**是等长的。

## 直接映射

一对一的映射关系，主存的一个块只能复制到Cache的一个特定行位置。  
Cache的行数 $j$ 和主存的块号 $i$

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

主存地址被划分为**块内地址**，**行地址**，**高位标记**三部分

根据Cache的行数将主存划分为若干各大组，每个大组包含 $2^c$ 个数据块，每个大组中地址偏移量相同的数据块，映射到Cache的同一行，该Cache行地址等于偏移量

优点：映射机制简单

缺点：机制不灵活，Cache命中率低

## 全相联映射

主存地址被划分为**块内地址**和**标记**两部分。将主存中一个块地址与块内容一起存入Cache行，其中块地址存放在标记部分。查找时主存地址中 $m$ 位块号和Cache中所有行的标记同时在比较器中进行比较。

优点：机制灵活，命中率高

缺点：硬件开销大

## 组相联映射

组相联映射方式将Cache的行分成 $2^{c-r}$ 组，每组 $2^r$ 行。主存字块存放到Cache中的哪一组是**固定的**，但是映射到组内的哪一行是**灵活的**。

主存中的低地址 $b$ 位对应块内地址，中间 $c - r$ 位对应Cache的组地址，高 $m - c + r$ 位对应行的标记。

**块内地址的大小**和Cache每行可以放多少个主存块有关

**组相联**，例如8路组相联代表的是每组有8行

## 查找过程

首先用中间 $c - r$ 位遭到Cache中的组，然后采用多路同时比较方式，将高 $m - c + r$ 位与该组 $2^r$ 行中所有标记同时进行比较，

## 优点

大大增加了映射的灵活性，同时硬件开销不大

## 替换策略

- 随机替换算法
- 先进先出算法
- 最近最少使用算法

## 写策略

### 写贯穿策略

当CPU写Cache命中时同时对主存执行写操作

当CPU写Cache不命中时

- 写不分配：直接修改主存，不将该数据所在的块复制到Cache行
- 写分配：先将数据所在块复制到Cache再进行Cache命中的操作

### 写回策略

相比写贯穿更快

当CPU写Cache命中时写操作只对Cache作用，而不修改主存

当CPU写Cache不命中时先将数据所在块复制到Cache再进行Cache命中的操作

需要为每个Cache分配一个修改位，被替换时写入主存。

## 主存储器

位于CPU外，用来存放**已被启动**的数据和**所有的**数据

用DRAM实现，速度较快，容量较大

## 特点

- 可以被CPU直接读取
- 一般由半导体材质构成
- 支持随机存取
- 与辅存相比速度快，价格高，容量小
- 一次从主存读出或写入的位数，可以**不等于**存储字的长度，也可以**不等于**编址单位

## 通信方式

### 异步方式

CPU发读命令，等待存储器发送“完成”信号，从数据线中读取数据

### 同步方式

CPU和主存在一个时钟信号控制下工作，不需要应答信号，CPU总是在发出读命令**后确定的几个时钟周期**后去取数据

## 辅助寄存器（外存）

用于存放暂时不用的数据和程序，属于永久存储器

## 磁盘存储器

分为软磁盘存储器和硬磁盘存储器

### 构成

- 磁盘控制器
- 磁盘驱动器
- 磁盘盘片

### 特点

既是微型计算机最重要的外部存储器，又是重要的输入输出设备

### 硬盘

大多采用**温盘**（温切斯特盘）

道密度：半径方向上每英寸磁道数量

位密度：磁道上每英寸包含的二进制位数

磁盘容量 = 盘片数量 × 2 × 磁道数量 × 每条磁道上的扇区数量 × 扇区容量

### 访问数据

寻道：将磁头移动到目标磁道上，所需的时间称为**寻道时间**

寻道成功后等待目标扇区旋转到磁头之下，称为**等待时间**或**旋转延迟**，通常取磁盘转动一周所需要的时间的一半作为**平均旋转延迟时间**

定位完成后传输数据的时间称为**传输时间**

## 光盘存储器

### 分类

- 只读光盘CD-ROM
- 一次性写入光盘WORM
- 可擦除光盘EOD

### 接口类型

- IDE
- EIDE

## 闪存盘

在EPROM基础上增加了电路的电擦除和重新编程功能

## 存储器容量拓展的三种方法

### 位扩展

位并联

芯片的数据线分别对应于所搭建的寄存器的高若干位和低若干位

### 字扩展

地址串联

用1K x 8位的SRAM存储器扩展为2K x 8位的SRAM存储器

### 字位扩展

结合位扩展和字扩展

## 高速存储器

弥补CPU与主存之间的速度差距

- 提高主存芯片的访问速度
  - 提高芯片本身的访问速度
  - 增加SRAM材质的行换成
- 采用并行技术
  - 双端口存储器（同时读写）
  - 多通道内存



- 多体交叉存储器
- 加入一个高速Cache

## 双端口存储器

特点：存储数据共享

## 多体交叉存储器

特点：通过更改主存的组织方式，在不改变存储器存取周期的情况下，提高存储器带宽

编址方法：交叉编址，任意两个相邻地址的物理单元不属于同一个存储体，同一个存储体内的地址是不连续的

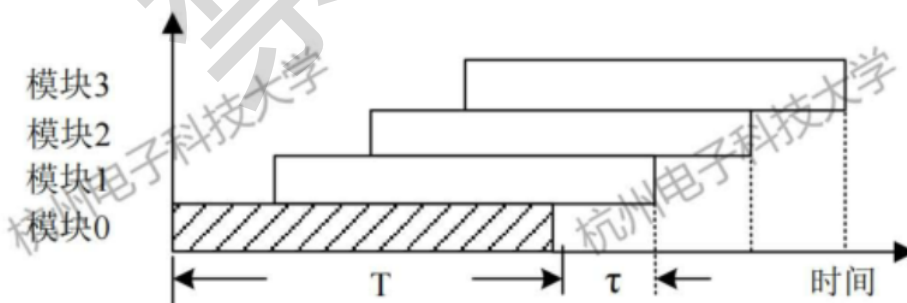
## 存储模块的启动方式

### 轮流启动

每个存储模块读写的数据宽度等于存储器总线中数据总线宽度时采用模块轮流启动方式

使用流水线方式存取满足  $T = M\tau$

- 总线传输速度为  $\tau$
- 存储器交叉编址模块数量为  $M$
- 存取模块存取一个字的周期为  $T$



### 同时启动

所有存储模块一次读写的数据宽度总和等于存储器总线中的数据总线宽度

## 相联存储器

一个字是通过它的部分内容而不是他的地址进行检索

禁止商用