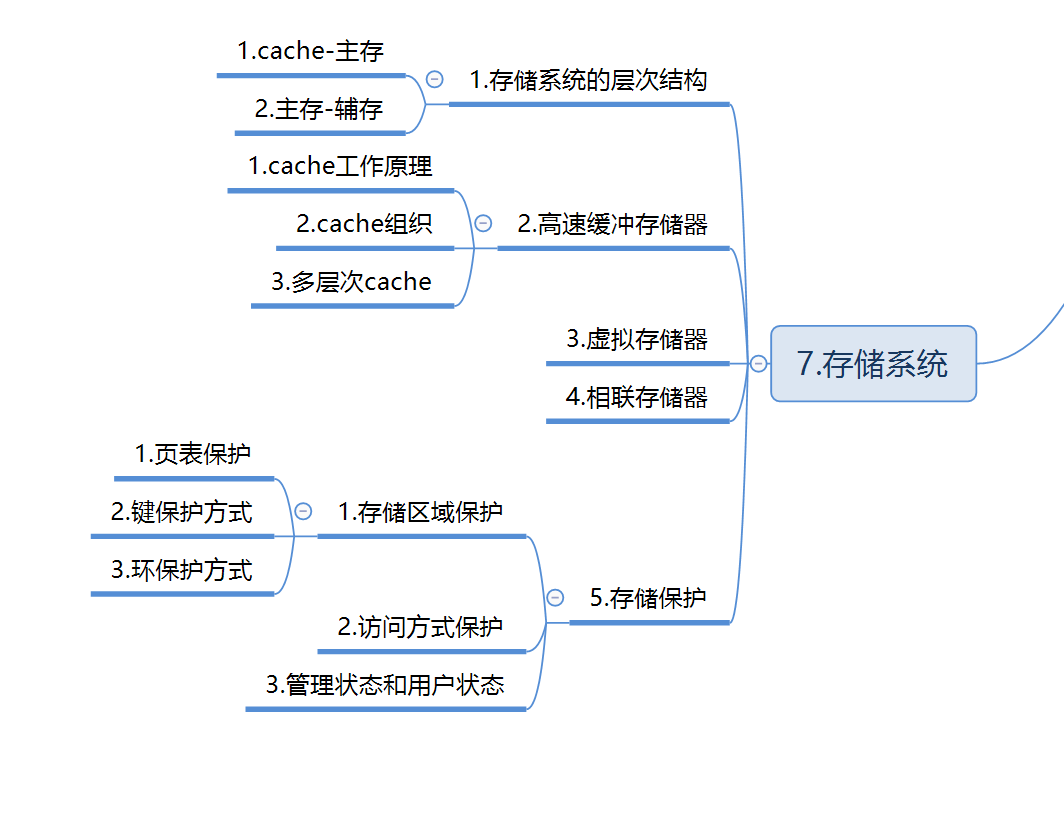
# 存储系统



## 存储系统的层次结构

1.1虚拟地址和物理地址

虚拟地址对应虚存容量（比主存实际空间大得多）

物理地址对应主存容量

1.2层次结构

层次结构分为2种：主存-辅存、cache-主存。构成三级存储层次：cache-主存-辅存。

## 高速缓冲存储器

* 1. cache工作原理

2.1.1局部性原理

对局部范围的存储器地址频繁访问的现象就称为程序访问的局部性，根据局部性原理，在主存和CPU之间设置了cache，也就是把程序中最近常用的部分驻留在cache中。

2.1.2主存和cache的基本结构

1. 主存

主存有2n个单元，地址码为n位，n=m+b：主存的块数M=2m块，块内的字节数（存储单元数即容量）B=2b

1. cache

cache地址码为（c+b），cache的块数为2c，块内字节数为2b与主存相同。cache中，每一块外加一个标记，指明它是主存的哪一块副本。

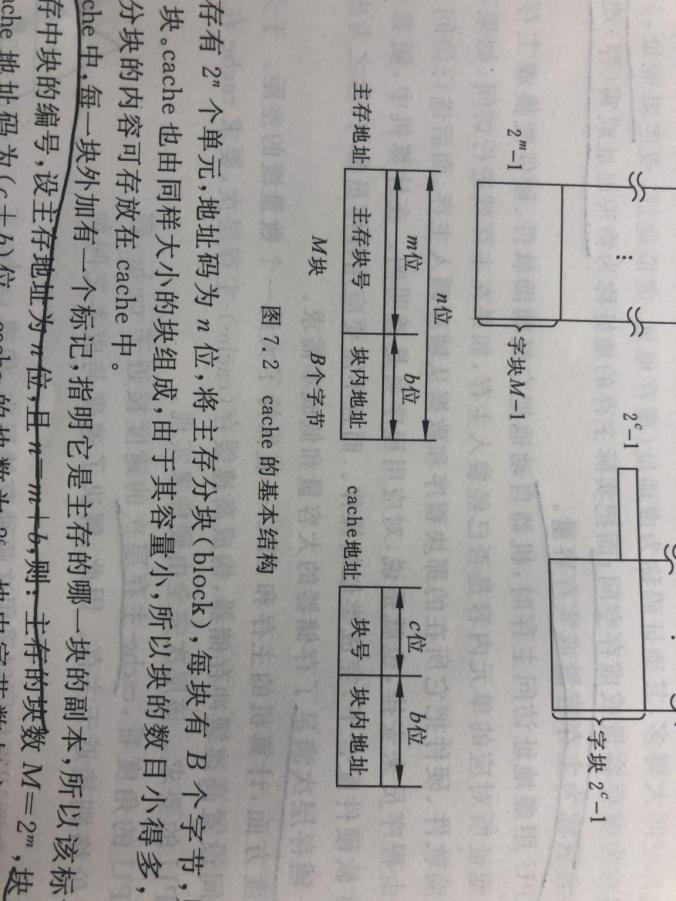


图 1主存和cacha的基本结构

2.1.3工作原理

1. 在CPU和cache之间通常一次传送一个字节
2. 命中率

命中率可以测量cache的效率，指CPU所要访问的信息是否在cache中的比率；失效率为索要访问的信息不在cache的比率。当cache的容量超过一定值后，命中率随容量的增加将不会有明显的增长。

1. 替换

主存和cache的替换策略由替换部件实现

1. cache和主存存储内容一致性

cache中保存的字块是主存中相应字块的一个副本。如果遇到写操作，就会遇到如何保持存储内容一致性问题，所以由3种写入方式。

（1）标志交换方式（写回法）：暂时只向cache写入，并用标志加以注明，直到经过修改的字块被从cache种替换出来时，才一次写入主存。特点：主存种的字块未经随时修改而可能失效。

（2）通过式写（写通法）：每次写入cache时也同时写入主存，使cache和主存保持一致。特点：能随时保持主存数据的正确性，但有可能增加多次不必要的向主存写入。

（3）当被修改的单元不在cache时，可以直接对主存进行，而不写入cache。

5）有效位

机器刚启动时，有效位置0；在程序执行过程中，cache不命中时将指令块或数据块从主存中调入，并将此块的有效位置1。

1. 平均存取时间=h\*tC+（1-h）（tC+tM）

h：命中率、tC：cache的存取时间、tM：主存的存取时间

* 1. cache组织

2.2.1地址映像

地址映像：用某种函数关系将主存地址映像到cache上。

主存被分为2m个块，字块大小为2b个字；cache被分为2c个同样大小的块，字块大小为2b个字。

1. 直接映像

特点：映像函数为j=i mod 2c（j是cache的字块号，i是主存的字块号）。主存的字块只能对应唯一的cache字块。

主存地址（m+b位）=主存地址块标记（t位）+cache字块地址（c位）+字块内地址（b位）

1. 全相联映像

特点：主存的字块可以映像到cache的任何一个字块上，也能从占满的cache中替换出任何一个旧字块。

主存地址（ m+b位）=主存地址块标记（m=t+c位）+字块内地址（b位）

1. 组相联映像

特点：把cache字块分为2c组，每组包含2r个字块。组间直接映像，组内为全相联映像

主存地址=主存地址块标记（区地址t+块地址r位）+组地址（c位）+字块内地址（b位）

2.2.2替换算法

先进先出（FIFO）算法和近期最少使用（LRU）算法

2.2.3 cache的地址监听

输入输出设备向存储器写入数据时，若该地址与cache相应单元的标记相符且有效位为1，此时会造成cache和存储器数据不一致性。

解决方案：将cache的有效位置0，这样当CPU再访问时，需要去存储器再取数据。

2.3多层次cache

2.3.1指令cache和数据cache

将指令cache和数据cache分开，组成为两个独立的cache，称为哈佛结构。

2.3.2多层次cache结构

二级cache方案，将一级cache（L1）和容量更大的第二级cache（L2）设置在cpu内部。Cpu访存时，如果L1不命中，则访问L2。

## 虚拟存储器

## 相联存储器

4.1简介

相联存储器不按地址访问存储器，而按所存数据字的全部内容或部分内容进行查找（或检索），比较操作是并行。

4.2相联存储器基本组成

CR比较数寄存器，存放要比较的数（要检索的内容）。

MR屏蔽寄存器，当按比较数的部分内容进行检索时，把MR中要比较的位置成1，不要比较的位置成0。

SRR查找结果寄存器，加入比较结果第i个字满足要求，则SRR中的第i位为1，其余各位为0。

WSR有字选择寄存器，用来确定哪些字参与检索。

## 存储保护

5.1存储区域保护

1. 页表保护

每个程序都有自己的页表和段表，无论地址如何出错，也只能影响到分配给该程序的主存页面。页表保护是在生成主存地址前的保护。但若在地址变换过程中出现错误，形成错误主存地址，则保护无效。

1. 键保护

为主存的每一页分配一个存储键，相当于一把锁；为每道程序分配一个访问键，相当于一把钥匙。当访问主存的某一页时，访问键要和存储键相比较，若两键相符，则允许访问该页。

1. 环保护方式

按系统程序和用户程序的重要性和对整个系统正常运行的影响程度进行分环，赋予环号。环号越大，等级越低。操作系统的环号小与用户程序的环号。

环号小的可以去环号大的，同级环号也可以；环号大的去环号小的，则产生中断，由操作系统处理。

5.2访问方式保护

访问方式保护有RWE这3种方式及其逻辑组合。

5.3管理状态和用户状态

计算机执行程序时把工作状态分为2种，一种是执行操作系统或管理程序时所处的状态，称为管理状态；另一种是执行用户程序时所处的状态，称为用户状态。

为防止用户编辑出错而影响整个系统工作，在机器中设置了一些特权指令（只有操作系统等系统程序才能使用），假如在用户程序中出现特权指令，则在执行到该指令时立即中止执行并发出中断。