

含 8 KB 的片内缓存。

可是,由于片内缓存在芯片内,其容量不可能很大,这就可能致使 CPU 欲访问的信息不在缓存内,势必通过系统总线访问主存,访问次数多了,整机速度就会下降。如果在主存与片内缓存之间再加一级缓存,称为片外缓存,由比主存动态 RAM 和 ROM 存取速度更快的静态 RAM 组成。而且不使用系统总线作为片外缓存与 CPU 之间的传送路径,使用一个独立的数据路径,以减轻系统总线的负担。那么,从片外缓存调入片内缓存的速度就能提高,而 CPU 占用系统总线的时间也就大大下降,整机工作速度有明显改进。这种由片外缓存和片内缓存组成的 Cache 称为两级缓存,并称片内缓存为第一级,片外缓存为第二级。随着芯片集成度的提高,已有一些处理器将第二级 Cache 结合到处理器芯片上,改善了性能。

(2) 统一缓存和分立缓存

统一缓存是指指令和数据都存放在同一缓存内的 Cache;分立缓存是指指令和数据分别存放在两个缓存中,一个称为指令 Cache,另一个称为数据 Cache。两种缓存的选用主要考虑如下两个因素。

其一,它与主存结构有关,如果计算机的主存是统一的(指令、数据存储在同一主存内),则相应的 Cache 采用统一缓存;如果主存采用指令、数据分开存储的方案,则相应的 Cache 采用分立缓存。

其二,它与机器对指令执行的控制方式有关。当采用超前控制或流水线控制方式时,一般都采用分立缓存。

所谓超前控制,是指在当前指令执行过程尚未结束时就提前将下一条准备执行的指令取出,称为超前取指或指令预取。所谓流水线控制实质上是多条指令同时执行(详见第8章),又可视为指令流水。当然,要实现同时执行多条指令,机器的指令译码电路和功能部件也需多个。超前控制和流水线控制特别强调指令的预取和指令的并行执行,因此,这类机器必须将指令 Cache 和数据 Cache 分开,否则可能出现取指和执行过程对统一缓存的争用。如果此刻采用统一缓存,则在执行部件向缓存发出取数请求时,一旦指令预取机构也向缓存发出取指请求,那么统一缓存只能先满足执行部件请求,将数据送到执行部件,而让取指请求暂时等待,显然达不到预取指令的目的,从而影响指令流水的实现。可见,这类机器将两种缓存分立尤为重要。

图 4.52 为 Pentium 4 处理器框图。

图中有两级共 3 个 Cache,其中一级 Cache 分 L1 指令 Cache 和 L1 数据 Cache,另外还有一个二级 L2 Cache。

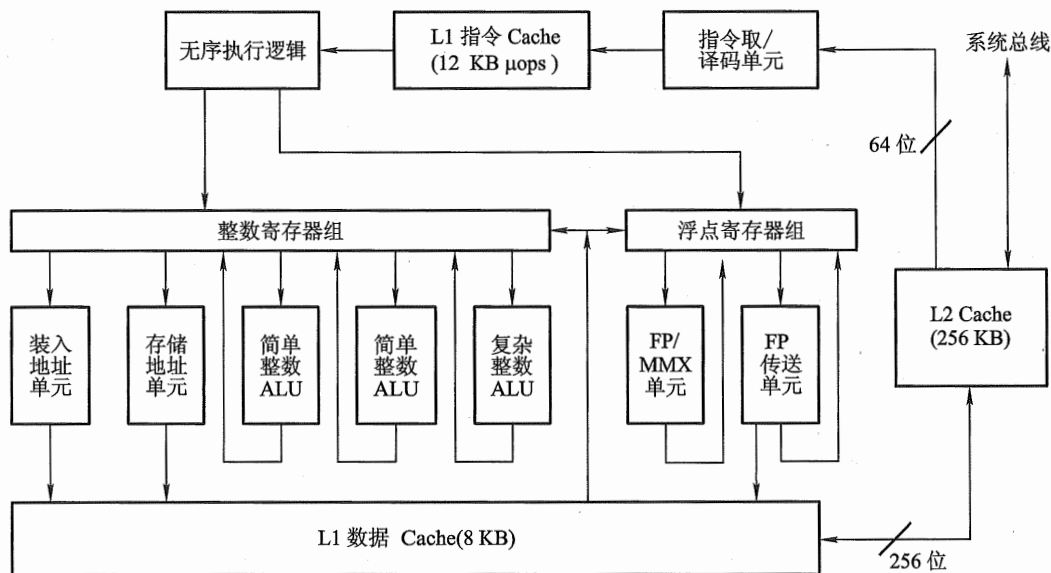


图 4.52 Pentium 4 处理器框图

图 4.53 是 PowerPC 620 处理器的示意图。

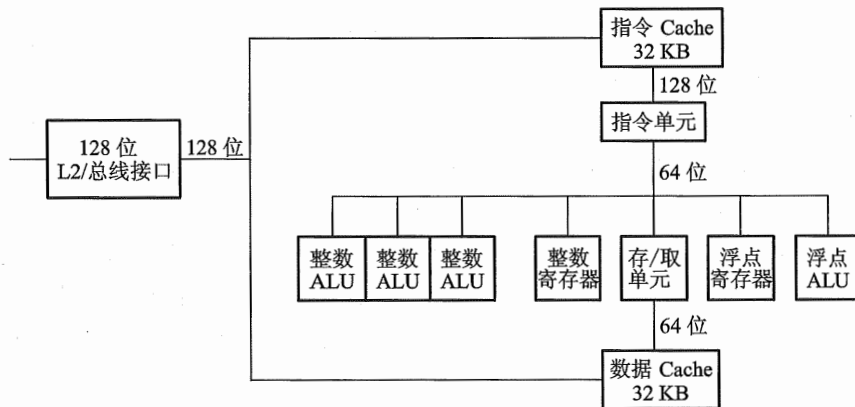


图 4.53 PowerPC 620 处理器框图

图中也有两个 Cache。数据 Cache 通过存/取单元支持整数和浮点操作；指令 Cache 为只读存储器，支持指令单元。执行部件是 3 个可并行操作的整数 ALU 和一个浮点运算部件（有独立的寄存器和乘、加、除部件）。

4.3.2 Cache—主存地址映射

由主存地址映射到 Cache 地址称为地址映射。地址映射方式很多,有直接映射(固定的映射关系)、全相联映射(灵活性大的映射关系)、组相联映射(上述两种映射的折中)。

1. 直接映射

图 4.54 示出了直接映射方式主存与缓存中字块的对应关系。

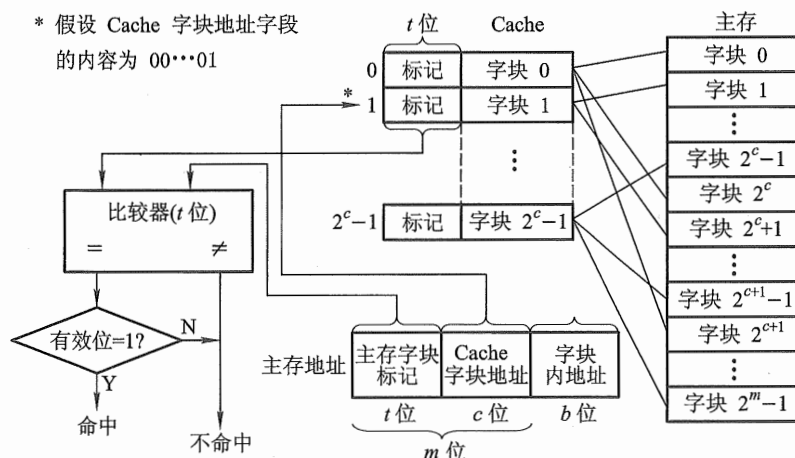


图 4.54 直接映射

图中每个主存块只与一个缓存块相对应,映射关系式为

$$i = j \bmod C \quad \text{或} \quad i = j \bmod 2^c$$

其中, i 为缓存块号, j 为主存块号, C 为缓存块数。映射结果表明每个缓存块对应若干个主存块,如表 4.4 所示。

表 4.4 直接映射方式主存块和缓存块的对应关系

缓存块	主存块
0	$0, C, \dots, 2^m - C$
1	$1, C + 1, \dots, 2^m - C + 1$
...	...
$C - 1$	$C - 1, 2C - 1, \dots, 2^m - 1$

这种方式的优点是实现简单,只需利用主存地址的某些位直接判断,即可确定所需字块是否在缓存中。由图 4.54 可见,主存地址高 m 位被分成两部分:低 c 位是指 Cache 的字块地址,高 t 位($t = m - c$)是指主存字块标记,它被记录在建立了对应关系的缓存块的“标记”位中。当缓存接

到 CPU 送来的主存地址后,只需根据中间 c 位字段(假设为 $00\cdots 01$)找到 Cache 字块 1,然后根据字块 1 的“标记”是否与主存地址的高 t 位相符来判断,若符合且有效位为“1”(有效位用来识别 Cache 存储块中的数据是否有效,因为有时 Cache 中的数据是无效的,例如,在初始时刻 Cache 应该是“空”的,其中的内容是无意义的),表示该 Cache 块已和主存的某块建立了对应关系(即已命中),则可根据 b 位地址从 Cache 中取得信息;若不符合,或有效位为“0”(即不命中),则从主存读入新的字块来替代旧的字块,同时将信息送往 CPU,并修改 Cache“标记”。如果原来有效位为“0”,还得将有效位置成“1”。

直接映射方式的缺点是不够灵活,因每个主存块只能固定地对应某个缓存块,即使缓存内还空着许多位置也不能占用,使缓存的存储空间得不到充分的利用。此外,如果程序恰好要重复访问对应同一缓存位置的不同主存块,就要不停地替换,从而降低命中率。

2. 全相联映射

全相联映射允许主存中每一字块映射到 Cache 中的任何一块位置上,如图 4.55 所示。这种映射方式可以从已被占满的 Cache 中替换出任一旧字块。显然,这种方式灵活,命中率也更高,缩小了块冲突率。与直接映射相比,它的主存字块标记从 t 位增加到 $t+c$ 位,这就使 Cache“标记”的位数增多,而且访问 Cache 时主存字块标记需要和 Cache 的全部“标记”位进行比较,才能判断出所访问主存地址的内容是否已在 Cache 内。这种比较通常采用“按内容寻址”的相联存储器(见附录 4A)来完成。

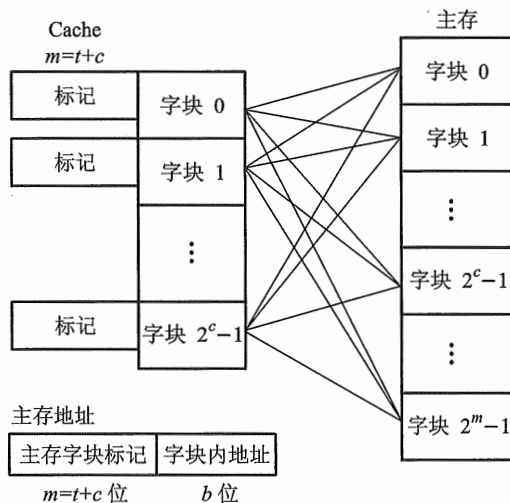


图 4.55 全相联映射

总之,这种方式所需的逻辑电路甚多,成本较高,实际的 Cache 还要采用各种措施来减少地址的比较次数。

3. 组相联映射

组相联映射是对直接映射和全相联映射的一种折中。它把 Cache 分为 Q 组, 每组有 R 块, 并有以下关系:

$$i = j \bmod Q$$

其中, i 为缓存的组号, j 为主存的块号。某一主存块按模 Q 将其映射到缓存的第 i 组内, 如图 4.56 所示。

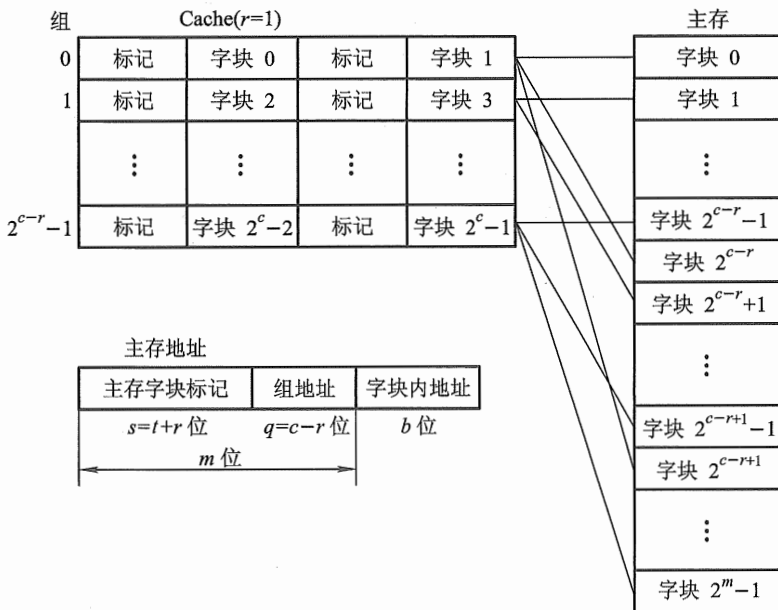


图 4.56 组相联映射

组相联映射的主存地址各段与直接映射 (参见图 4.54) 相比, 还是有区别的。图 4.54 中 Cache 字块地址字段由 c 位变为组地址字段 q 位, 且 $q=c-r$, 其中 2^c 表示 Cache 的总块数, 2^q 表示 Cache 的分组个数, 2^r 表示组内包含的块数。主存字块标记字段由 t 位变为 $s=t+r$ 位。为了便于理解, 假设 $c=5, q=4$, 则 $r=c-q=1$ 。其实际含义为: Cache 共有 $2^c=32$ 个字块, 共分为 $2^q=16$ 组, 每组内包含 $2^r=2$ 块。组内 2 块的组相联映射又称为二路组相联。

根据上述假设条件, 组相联映射的含义是: 主存的某一字块可以按模 16 映射到 Cache 某组的任一字块中。即主存的第 0, 16, 32... 字块可以映射到 Cache 第 0 组 2 个字块中的任一字块; 主存的第 15, 31, 47... 字块可以映射到 Cache 第 15 组中的任一字块。显然, 主存的第 j 块会映射到 Cache 的第 i 组内, 两者之间一一对应, 属直接映射关系; 另一方面, 主存的第 j 块可以映射到 Cache 的第 i 组内中的任一块, 这又体现出全相联映射关系。可见, 组相联映射的性能及其复杂性介于直接映射和全相联映射两者之间, 当 $r=0$ 时是直接映射方式, 当 $r=c$ 时是全相联映射方式。

例 4.8 假设主存容量为 512 KB,Cache 容量为 4 KB,每个字块为 16 个字,每个字 32 位。

- (1) Cache 地址有多少位? 可容纳多少块?
- (2) 主存地址有多少位? 可容纳多少块?
- (3) 在直接映射方式下,主存的第几块映射到 Cache 中的第 5 块(设起始字块为第 1 块)?
- (4) 画出直接映射方式下主存地址字段中各段的位数。

解:(1) 根据 Cache 容量为 4 KB($2^{12}=4\text{ K}$),Cache 地址为 12 位。由于每字 32 位,则 Cache 共有 $4\text{ KB}/4\text{ B}=1\text{ K}$ 字。因每个字块 16 个字,故 Cache 中有 $1\text{ K}/16=64$ 块。

(2) 根据主存容量为 512 KB($2^{19}=512\text{ K}$),主存地址为 19 位。由于每字 32 位,则主存共有 $512\text{ KB}/4\text{ B}=128\text{ K}$ 字。因每个字块 16 个字,故主存中共 $128\text{ K}/16=8\text{ 192}$ 块。

(3) 在直接映射方式下,由于 Cache 共有 64 块,主存共有 8 192 块,因此主存的 $5, 64+5, 2\times 64+5, \dots, 2^{13}-64+5$ 块能映射到 Cache 的第 5 块中。

(4) 在直接映射方式下,主存地址字段的各段位数分配如图 4.57 所示。其中字块内地址为 6 位(4 位表示 16 个字,2 位表示每字 32 位),缓存共 64 块,故缓存字块地址为 6 位,主存字块标记为主存地址长度与 Cache 地址长度之差,即 $19-12=7$ 位。

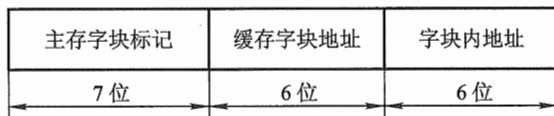


图 4.57 例 4.8 主存地址各字段的分配

例 4.9 假设主存容量为 512 K×16 位,Cache 容量为 4 096×16 位,块长为 4 个 16 位的字,访存地址为字地址。

- (1) 在直接映射方式下,设计主存的地址格式。
- (2) 在全相联映射方式下,设计主存的地址格式。
- (3) 在二路组相联映射方式下,设计主存的地址格式。

(4) 若主存容量为 512 K×32 位,块长不变,在四路组相联映射方式下,设计主存的地址格式。

解:(1) 根据 Cache 容量为 $4\text{ 096}=2^{12}$ 字,得 Cache 字地址为 12 位。根据块长为 4,且访存地址为字地址,得字块内地址为 2 位,即 $b=2$,且 Cache 共有 $4\text{ 096}/4=1\text{ 024}=2^{10}$ 块,即 $c=10$ 。根据主存容量为 $512\text{ K}=2^{19}$ 字,得主存字地址为 19 位。在直接映射方式下,主存字块标记为 $19-12=7$ 。主存的地址格式如图 4.58(a) 所示。

(2) 在全相联映射方式下,主存字块标记为 $19-b=19-2=17$ 位,其地址格式如图 4.58(b) 所示。

(3) 根据二路组相联的条件,一组内有 2 块,得 Cache 共分 $1\text{ 024}/2=512=2^9$ 组,即 $q=9$,主存字块标记为 $19-q-b=19-9-2=8$ 位,其地址格式如图 4.58(c) 所示。

(4) 若主存容量改为 $512\text{ K}\times 32$ 位,即双字宽存储器,块长仍为 4 个 16 位的字,访存地址仍为字地址,则主存容量可写为 $1\,024\text{ K}\times 16$ 位,得主存地址为 20 位。由四路组相联,得 Cache 共分 $1\,024/4=256=2^q$ 组,即 $q=8$ 。对应该条件下,主存字块标记为 $20-8-2=10$ 位,其地址格式如图 4.58(d) 所示。

主存字块标记	Cache 字块地址	字块内地址
7	10	2

(a) 直接映射方式主存地址格式

主存字块标记	字块内地址
17	2

(b) 全相联映射方式主存地址格式

主存字块标记	组地址	字块内地址
8	9	2

(c) 二路组相联映射方式主存地址格式

主存字块标记	组地址	字块内地址
10	8	2

(d) 四路组相联映射方式双字宽主存地址格式

图 4.58 例 4.9 主存地址格式

例 4.10 假设 Cache 的工作速度是主存的 5 倍,且 Cache 被访问命中的概率为 95%,则采用 Cache 后,存储器性能提高多少?

解: 设 Cache 的存取周期为 t ,主存的存取周期为 $5t$,则系统的平均访问时间为

$$t_a = 0.95 \times t + 0.05 \times 5t = 1.2t$$

性能为原来的 $5t/1.2t=4.17$ 倍,即提高了 3.17 倍。

例 4.11 设某机主存容量为 16 MB,Cache 的容量为 8 KB。每字块有 8 个字,每字 32 位。设计一个四路组相联映射的 Cache 组织。

(1) 画出主存地址字段中各段的位数。

(2) 设 Cache 初态为空,CPU 依次从主存第 0,1,2,⋯,99 号单元读出 100 个字(主存一次读出一个字),并重复此次序读 10 次,问命中率是多少?

(3) 若 Cache 的速度是主存速度的 5 倍,试问有 Cache 和无 Cache 相比,速度提高多少倍?

(4) 系统的效率为多少?

解:(1) 根据每个字块有 8 个字,每个字 32 位,得出主存地址字段中字块内地址字段为 5 位,其中 3 位为字地址,2 位为字节地址。

根据 Cache 容量为 $8\text{ KB} = 2^{13}\text{ B}$,字块大小为 2^5 B ,得 Cache 共有 2^8 块,故 $c = 8$ 。根据四路组相联映射 $2^r = 4$,得 $r = 2$,则 $q = c - r = 8 - 2 = 6$ 位。

根据主存容量为 $16\text{ MB} = 2^{24}\text{ B}$,得出主存地址字段中主存字块标记为 $24 - 6 - 5 = 13$ 位。

主存地址字段各段格式如图 4.59 所示。

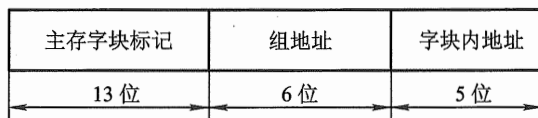


图 4.59 例 4.10 主存地址字段

(2) 由于每个字块中有 8 个字,而且初态 Cache 为空,因此 CPU 读第 0 号单元时,未命中,必须访问主存,同时将该字所在的主存块调入 Cache 第 0 组中的任一块内,接着 CPU 读 1~7 号单元时均命中。同理,CPU 读第 8,16,⋯,96 号单元时均未命中。可见 CPU 在连续读 100 个字中共有 13 次未命中,而后 9 次循环读 100 个字全部命中,命中率为

$$\frac{100 \times 10 - 13}{100 \times 10} = 0.987$$

(3) 根据题意,设主存取周期为 $5t$,Cache 的存取周期为 t ,没有 Cache 的访问时间为 $5t \times 1\,000$,有 Cache 的访问时间为 $t(1\,000 - 13) + 5t \times 13$,则有 Cache 和没有 Cache 相比,速度提高的倍数为

$$\frac{5t \times 1\,000}{t(1\,000 - 13) + 5t \times 13} - 1 \approx 3.75$$

(4) 根据(2)求得的命中率 0.987,主存的存取周期为 $5t$,Cache 的存取周期为 t ,得系统的效率为

$$\frac{t}{0.987 \times t + (1 - 0.987) \times 5t} \times 100\% = 95\%$$

4.3.3 替换策略

当新的主存块需要调入 Cache 并且它的可用空间位置又被占满时,需要替换掉 Cache 的数据,这就产生了替换策略(算法)问题。在直接映射的 Cache 中,由于某个主存块只与一个 Cache 字块有映射关系,因此替换策略很简单。而在组相联和全相联映射的 Cache 中,主存块可以写入

Cache 中若干位置,这就有一个选择替换掉哪一个 Cache 字块的问题,即所谓替换算法问题。理想的替换方法是把未来很少用到的或者很久才用到的数据块替换出来,但实际上很难做到。常用的替换算法有先进先出算法、近期最少使用算法和随机法。

1. 先进先出(First-In-First-Out, FIFO)算法

FIFO 算法选择最早调入 Cache 的字块进行替换,它不需要记录各字块的使用情况,比较容易实现,开销小,但没有根据访存的局部性原理,故不能提高 Cache 的命中率。因为最早调入的信息可能以后还要用到,或者经常要用到,如循环程序。

2. 近期最少使用(Least Recently Used, LRU)算法

LRU 算法比较好地利用访存局部性原理,替换出近期用得最少的字块。它需要随时记录 Cache 中各字块的使用情况,以便确定哪个字块是近期最少使用的字块。它实际是一种推测的方法,比较复杂,一般采用简化的方法,只记录每个块最近一次使用的时间。LRU 算法的平均命中率比 FIFO 的高。

3. 随机法

随机法是随机地确定被替换的块,比较简单,可采用一个随机数产生器产生一个随机的被替换的块,但它也没有根据访存的局部性原理,故不能提高 Cache 的命中率。

4.4 辅助存储器

4.4.1 概述

1. 辅助存储器的特点

辅助存储器作为主存的后援设备又称为外部存储器,简称外存,它与主存一起组成了存储器系统的主存-辅存层次。与主存相比,辅存具有容量大、速度慢、价格低、可脱机保存信息等特点,属“非易失性”存储器。而主存具有速度快、成本高、容量小等特点,而且大多由半导体芯片构成,所存信息无法永久保存,属“易失性”存储器。

目前,广泛用于计算机系统的辅助存储器有硬磁盘、软磁盘、磁带、光盘等。前三种均属磁表面存储器。

磁表面存储器是在不同形状(如盘状、带状等)的载体上涂有磁性材料层,工作时,靠载磁体高速运动,由磁头在磁层上进行读/写操作,信息被记录在磁层上,这些信息的轨迹就是磁道。磁盘的磁道是一个个同心圆,如图 4.60(a)所示,磁带的磁道是沿磁带长度方向的直线,如图 4.60(b)所示。

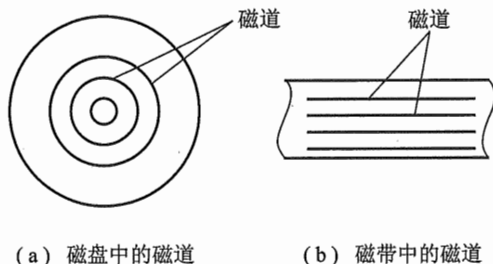


图 4.60 磁盘和磁带的磁道示意图

2. 磁表面存储器的主要技术指标

(1) 记录密度

记录密度通常是指单位长度内所存储的二进制信息量。磁盘存储器用道密度和位密度表示；磁带存储器则用位密度表示。磁盘沿半径方向单位长度的磁道数为道密度，单位是 tpi (Track Per Inch, 道每英寸) 或 tpm (道每毫米)。为了避免干扰，磁道与磁道之间需保持一定距离，相邻两条磁道中心线之间的距离称为道距，因此道密度 D_t 等于道距 P 的倒数，即

$$D_t = \frac{1}{P}$$

单位长度磁道能记录二进制信息的位数，称为位密度或线密度，单位是 bpi (Bits Per Inch, 位每英寸) 或 bpm (位每毫米)。磁带存储器主要用位密度来衡量，常用的磁带有 800 bpi、1 600 bpi、6 250 bpi 等。对于磁盘，位密度 D_b 可按下式计算：

$$D_b = \frac{f_t}{\pi d_{\min}}$$

其中， f_t 为每道总位数， d_{\min} 为同心圆中的最小直径。

在磁盘各磁道上所记录的信息量是相同的，而位密度不同，一般泛指磁盘位密度时，是指最内圈磁道上的位密度（最大位密度）。

(2) 存储容量

存储容量是指外存所能存储的二进制信息总数量，一般以位或字节为单位。以磁盘存储器为例，存储容量可按式计算：

$$C = n \times k \times s$$

其中， C 为存储总容量， n 为存放信息的盘面数， k 为每个盘面的磁道数， s 为每条磁道上记录的二进制代码数。

磁盘有格式化容量和非格式化容量两个指标。非格式化容量是磁表面可以利用的磁化单元总数。格式化容量是指按某种特定的记录格式所能存储信息的总量，即用户可以使用的容量，它

一般为非格式化容量的 60%~70%。

(3) 平均寻址时间

由存取方式分类可知,磁盘采取直接存取方式,寻址时间分为两个部分,其一是磁头寻找目标磁道的找道时间 t_s ,其二是找到磁道后,磁头等待欲读/写的磁道区段旋转到磁头下方所需要的等待时间 t_w 。由于从最外圈磁道找到最里圈磁道和寻找相邻磁道所需时间是不等的,而且磁头等待不同区段所花的时间也不等,因此,取其平均值,称为平均寻址时间 T_a ,它是平均找道时间 t_{sa} 和平均等待时间 t_{wa} 之和:

$$T_a = t_{sa} + t_{wa} = \frac{t_{smax} + t_{smin}}{2} + \frac{t_{wmax} + t_{wmin}}{2}$$

平均寻址时间是磁盘存储器的一个重要指标。硬磁盘的平均寻址时间比软磁盘的平均寻址时间短,所以硬磁盘存储器比软磁盘存储器速度快。

磁带存储器采取顺序存取方式,磁头不动,磁带移动,不需要寻找磁道,但要考虑磁头寻找记录区段的等待时间,所以磁带寻址时间是指磁带空转到磁头应访问的记录区段所在位置的时间。

(4) 数据传输率

数据传输率 D_r 是指单位时间内磁表面存储器向主机传送数据的位数或字节数,它与记录密度 D_b 和记录介质的运动速度 V 有关:

$$D_r = D_b \times V$$

此外,辅存和主机的接口逻辑应有足够快的传送速度,用来完成接收/发送信息,以便主机与辅存之间正确无误地传送。

(5) 误码率

误码率是衡量磁表面存储器出错概率的参数,它等于从辅存读出时,出错信息位数和读出信息的总位数之比。为了减少出错率,磁表面存储器通常采用循环冗余码来发现并纠正错误。

4.4.2 磁记录原理和记录方式

1. 磁记录原理

磁表面存储器通过磁头和记录介质的相对运动完成读/写操作。

写入过程如图 4.61 所示。

写入时,记录介质在磁头下方匀速通过,根据写入代码的要求,对写入线圈输入一定方向和大小的电流,使磁头导磁体磁化,产生一定方向和强度的磁场。由于磁头与磁层表面间距非常小,磁力线直接穿透磁层表面,将对应磁头下方的微小区域磁化(称为磁化单元)。可以根据写入驱动电流的不同方向,使磁层表面被磁化的极性方向不同,以区别记录“0”或“1”。

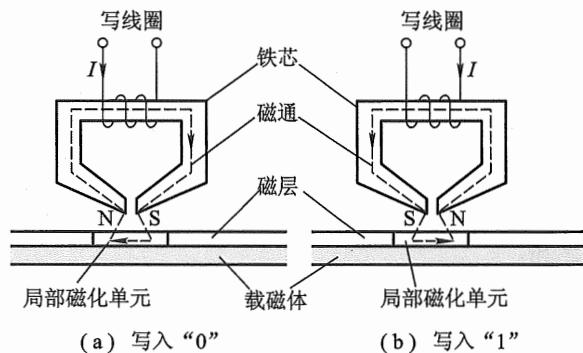


图 4.61 磁表面存储器写入原理

读出时,记录介质在磁头下方匀速通过,磁头相对于一个个被读出的磁化单元作切割磁力线的运动,从而在磁头读线圈中产生感应电势 e ,且 $e = -n \frac{d\phi}{dt}$ (n 为读出线圈匝数),其方向正好和磁通的变化方向相反。由于原来磁化单元的剩磁通 ϕ 的方向不同,感应电势方向也不同,便可读出“1”或“0”两种不同信息,如图 4.62 所示。

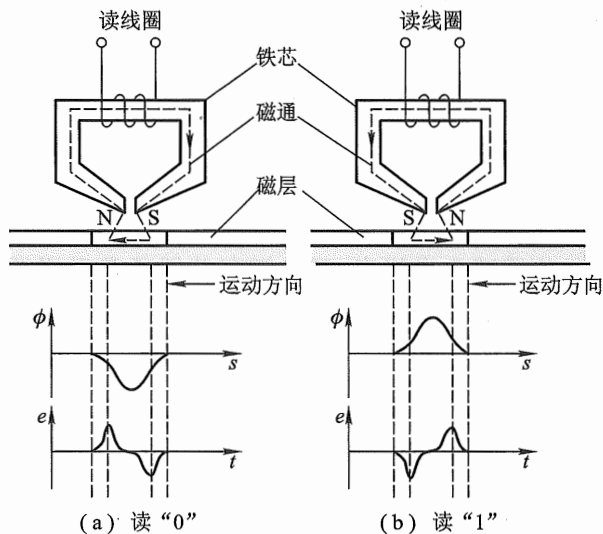


图 4.62 磁表面存储器读出原理

2. 磁表面存储器的记录方式

磁记录方式又称为编码方式,它是按某种规律将一串二进制数字信息变换成磁表面相应的磁化状态。磁记录方式对记录密度和可靠性都有很大影响,常用的记录方式有六种,如图 4.63 所示。

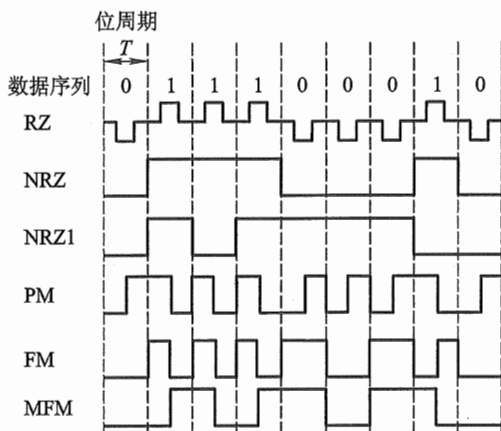


图 4.63 六种磁记录方式的写入电流波形

图中波形既代表了磁头线圈中的写入电流波形,也代表磁层上相应位置所记录的理想的磁通变化状态。

(1) 归零制(RZ)

归零制记录“1”时,通以正向脉冲电流,记录“0”时,通以反向脉冲电流。这样使其在磁表面形成两个不同极性的磁饱和状态,分别表示“1”和“0”。由于两位信息之间驱动电流归零,故称为归零制记录方式。这种方式在写入信息时很难覆盖原来的磁化区域,所以为了重新写入信息,在写入前,必须先抹去原存信息。这种记录方式原理简单,实施方便,但由于两个脉冲之间有一段间隔没有电流,相应的该段磁介质未被磁化,即该段空白,故记录密度不高,目前很少使用。

(2) 不归零制(NRZ)

不归零制记录信息时,磁头线圈始终有驱动电流,不是正向,便是反向,不存在无电流状态。这样,磁表面层不是正向被磁化,就是反向被磁化。当连续记录“1”或“0”时,其写电流方向不变,只有当相邻两信息代码不同时,写电流才改变方向,故称为“见变就翻”的不归零制。

(3) “见1就翻”的不归零制(NRZI)

“见1就翻”的不归零制在记录信息时,磁头线圈也始终有电流。但只有在记录“1”时电流改变方向,使磁层磁化方向发生翻转;记录“0”时,电流方向保持不变,使磁层的磁化方向也维持原来状态,因此称为“见1就翻”的不归零制。

(4) 调相制(PM)

调相制又称为相位编码(PE),其特点是记录“1”或“0”的相位相反。如:记录“0”时,写电流由负变正;记录“1”时,写电流由正变负(也可以相反定义),而且电流变化出现在一位信息记录时间的中间时刻,它以相位差为 180° 的磁化翻转方向来表示“1”和“0”。因此,当连续记录相同

信息时,在每两个相同信息的交界处,电流方向都要变化一次;若相邻信息不同,则两个信息位的交界处电流方向维持不变。调相制在磁带存储器中用得较多。

(5) 调频制(FM)

调频制的记录规则是:以驱动电流变化的频率不同来区别记录“1”还是“0”。当记录“0”时,在一位信息的记录时间内电流保持不变;当记录“1”时,在一位信息记录时间的中间时刻,使电流改变一次方向。而且无论记录“0”还是“1”,在相邻信息的交界处,线圈电流均变化一次。因此,写“1”时,在位单元的起始和中间位置都有磁通翻转;在写“0”时,仅在位单元起始位置有翻转。显然,记录“1”的磁翻转频率为记录“0”的两倍,故又称为倍频制。调频制记录方式被广泛应用在硬磁盘和软磁盘中。

(6) 改进型调频制(MFM)

这种记录方式基本上同调频制,即记录“0”时,在位记录时间内电流不变;记录“1”时,在位记录时间的中间时刻电流发生一次变化。两者不同之处在于,改进型调频制只有当连续记录两个或两个以上的“0”时,才在每位的起始处改变一次电流,不必在每个位起始处都改变电流方向。由于这一特点,在写入同样数据序列时,MFM比FM磁翻转次数少,在相同长度的磁层上可记录的信息量将会增加,从而提高了磁记录密度。FM制记录一位二进制代码最多是两次磁翻转,MFM制最多只要一次翻转,记录密度提高了一倍,故又称为倍密度记录方式。倍密度软磁盘即采用MFM记录方式。

此外还有一种二次改进的调频制(M^2FM),它是在MFM基础上改进的,其记录规则是:当连续记录“0”时,仅在第1个位起始处改变电流方向,以后的位交界处电流方向不变。

3. 评价记录方式的主要指标

评价一种记录方式的优劣标准主要反映在编码效率和自同步能力等方面。

(1) 编码效率

编码效率是指位密度与磁化翻转密度的比值,可用记录一位信息的最大磁化翻转次数来表示。例如,FM、PM记录方式中,记录一位信息最大磁化翻转次数为2,因此编码效率为50%;而MFM、NRZ、NRZ1三种记录方式的编码效率为100%,因为它们记录一位信息磁化翻转最多一次。

(2) 自同步能力

自同步能力是指从单个磁道读出的脉冲序列中所提取同步时钟脉冲的难易程度。从磁表面存储器的读出可知,为了将数据信息分离出来,必须有时间基准信号,称为同步信号。同步信号可以从专门设置用来记录同步信号的磁道中取得,这种方法称为外同步,如NRZ1制。图4.64画出了NRZ1制驱动电流、记录磁通、感应电势、同步脉冲、读出代码等几种波形的理想对应关系(图中未反映磁通变化的滞后现象)。读出时将读线圈获得的感应信号放大(负波还要反相)、整形,这样,对于每个记录的“1”都会得到一个正脉冲,再将它们与同步脉冲相“与”,即可得读出代码波形。

对于高密度的记录系统,可直接从磁盘读出的信号中提取同步信号,这种方法称为自同步。

自同步能力可用最小磁化翻转间隔和最大磁化翻转间隔之比值 R 来衡量。 R 越大,自同步能力也越强。例如,NRZ 和 NRZ1 方式在连续记录“0”时,磁层都不发生磁化磁转,而 NRZ 方式在连续记录“1”时,磁层也不发生磁化翻转,因此,NRZ 和 NRZ1 都没有自同步能力。而 PM、FM、MFM 记录方式均有自同步能力。FM 记录方式的最大磁化翻转间隔是 T (T 为一位信息的记录时间),最小磁化翻转间隔是 $T/2$,所以 $R_{FM} = 0.5$ 。

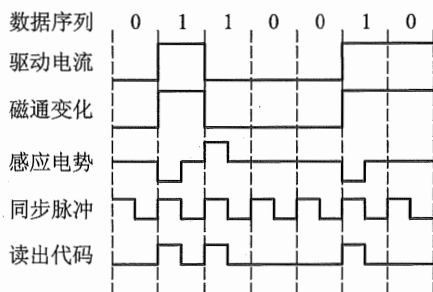


图 4.64 NRZ1 的读出代码波形

影响记录方式的优劣因素还有很多,如读分辨力、信息独立性(即某一位信息读出时出现误码而不影响后续其他信息位的正确性)、频带宽度、抗干扰能力以及实现电路的复杂性等。

除上述所介绍的 6 种记录方式外,还有成组编码记录方式,如 GCR(5.4) 编码,它广泛用于磁带存储器,游程长度受限码(RLL 码)是近年发展起来的用于高密度磁盘上的一种记录方式,在此均不详述。

4.4.3 硬磁盘存储器

硬磁盘存储器是计算机系统最主要的外存设备。第一个商品化的硬磁盘是由美国 IBM 公司于 1956 年研制而成的。60 多年来,无论在结构还是在性能方面,磁盘存储器有了很大的发展和改进。

1. 硬磁盘存储器类型

硬磁盘存储器的盘片是由硬质铝合金材料制成的,其表面涂有一层可被磁化的硬磁特性材料。按磁头的工作方式可分为固定磁头磁盘存储器和移动磁头磁盘存储器;按磁盘是否具有可换性又可分为可换盘磁盘存储器和固定盘磁盘存储器。

固定磁头的磁盘存储器,其磁头位置固定不动,磁盘上的每一个磁道都对应一个磁头,如图 4.65(a)所示,盘片也不可更换。其特点是省去了磁头沿盘片径向运动所需寻找磁道的时间,存取速度快,只要磁头进入工作状态即可进行读写操作。

移动磁头的磁盘存储器在存取数据时,磁头在盘面上做径向运动,这类存储器可以由一个盘片组成,如图 4.65(b)所示。也可由多个盘片装在一个同心主轴上,每个记录面各有一个磁头,如图 4.65(c)所示。

图 4.65(c)中含有 6 个盘片,除上下两外侧为保护面外,共有 10 个盘面可作为记录面,并对应 10 个磁头(有的磁盘组最外两侧盘面也可作为记录面,并分别与一个磁头对应)。所有这些磁头连成一体,固定在一个支架上可以移动,任何时刻各磁头都位于距圆心相等距离的磁道上,这组磁道称为一个柱面。目前,这类结构的硬磁盘存储器应用最广泛。最