

**01** 例题·cache 中的位数

假设一个直接映射的 cache，有 16KiB 的数据，块大小为 4 个字，地址为 32 位，那么该 cache 总共需要多少位？

**01** 答案

我们知道 16KiB 是  $4096(2^{12})$  个字，块大小是 4 个字 ( $2^2$ )，那么就有 1024 ( $2^{10}$ ) 个块。每个块有  $4 \times 32$  即 128 位的数据，加上  $32 - 10 - 2 - 2$  位的标记域，再加上一个有效位，因此，总的 cache 大小是

$$2^{10} \times (4 \times 32 + (32 - 10 - 2 - 2) + 1) = 2^{10} \times 147 = 147\text{Kib}$$

即能装 16KiB 数据的 cache 总共需要 18.4KiB 的容量。这个 cache 的总位数是数据存储量的 1.15 倍。 □

**01** 例题·将一个地址映射到多字大小的 cache 块中

考虑一个 cache 中有 64 个块，每块大小为 16 字节，那么字节地址 1200 将被映射到 cache 中的哪一块？

**01** 答案

块由下面公式给出：

$$(\text{块地址}) \bmod (\text{cache 中的块数})$$

其中块地址为

$$\text{字节地址} / \text{每块字节数}$$

注意，这个块地址包含了所有在

$$\text{字节地址} / \text{每块字节数} \times \text{每块字节数}$$

和

$$\text{字节地址} / \text{每块字节数} \times \text{每块字节数} + (\text{每块字节数} - 1)$$

之间的地址。

因此，由于每个块有 16 字节，字节地址 1200 对应的块地址为

$$1200/16 = 75$$

对应于 cache 中的块号  $(75 \bmod 64) = 11$ 。事实上，地址 1200 和 1215 之间的所有地址都映射在这一块。 □

较大的 cache 块能更好地利用空间局部性以降低缺失率。如图 5-11 所示，增加块大小通常会引起缺失率下降。而当块大小在 cache 容量中所占比例增加到一定程度时，缺失率也随之增加。这是因为此时 cache 中块的数量变得很少，对于这些块将会有大量的竞争发生。结果，就造成一个块中的数据在被多次访问之前就被替换出 cache。另一方面，对于一个太大的块，块中各个字之间的空间局部性也会降低，缺失率降低所带来的益处也会相应减少。

391

仅仅增加块大小所带来的一个更加严重的后果是缺失成本的增加。由较低存储器层次取出块并存放至 cache 中所花费的时间决定了缺失代价。取出块的时间可以分为两部分：第一个字的延迟时间和剩余部分块的传输时间。很显然，除非改变存储系统，否则，传输时间，也就是缺失代价将随着块大小的增大而增加。此外，当块越来越大时，缺失率的改善也开始降低。而当块过于大时，缺失代价的增长超过了缺失率的降低，因此 cache 的性能也随之降低。当然，如果把存储器设计得能更有效地传输较大的块时，我们就能增加块的大小并且进一步改善 cache 性能。这一点我们将在下一节讨论。