


基本参数	
参数	描述
$S=2^s$	组数
$E$	每个组的行数
$B=2^b$	块大小(字节)
$m=\log_2(M)$	(主存)物理地址位数

衍生出来的量	
参数	描述
$M=2^m$	内存地址的最大数量
$s=\log_2(S)$	组索引位数
$b=\log_2(B)$	块偏移位数
$t=m-(s+b)$	标记位数
$C=B \times E \times S$	不包括像有效位和标记位这样开销的高速缓存大小(字节)

图 6-26 高速缓存参数小结

 **练习题 6.9** 下表给出了几个不同的高速缓存的参数。确定每个高速缓存的高速缓存组数( $S$ )、标记位数( $t$ )、组索引位数( $s$ )以及块偏移位数( $b$ )。

高速缓存	$m$	$C$	$B$	$E$	$S$	$t$	$s$	$b$
1.	32	1024	4	1				
2.	32	1024	8	4				
3.	32	1024	32	32				

### 6.4.2 直接映射高速缓存

根据每个组的高速缓存行数  $E$ ，高速缓存被分为不同的类。每个组只有一行( $E=1$ )的高速缓存称为直接映射高速缓存(direct-mapped cache)(见图 6-27)。直接映射高速缓存是最容易实现和理解的，所以我们会以它为例来说明一些高速缓存工作方式的通用概念。

图 6-27 直接映射高速缓存( $E=1$ )。每个组只有一行

假设我们有这样一个系统，它有一个 CPU、一个寄存器文件、一个 L1 高速缓存和一个主存。当 CPU 执行一条读内存字  $w$  的指令，它向 L1 高速缓存请求这个字。如果 L1 高速缓存有  $w$  的一个缓存的副本，那么就得到 L1 高速缓存命中，高速缓存会很快抽取出自  $w$ ，并将它返回给 CPU。否则就是缓存不命中，当 L1 高速缓存向主存请求包含  $w$  的块的一个副本时，CPU 必须等待。当被请求的块最终从内存到达时，L1 高速缓存将这个块存放在它的一个高速缓存行里，从被存储的块中抽取出自  $w$ ，然后将它返回给 CPU。高速