

86. 以不同位作索引, 组别之间在内存中的映射关系不同, 组之间的距离也不同。若用高位索引, 连续的内存映射了同一个组, 访问时就会发生多次缓存不命中; 若采用中间位索引, 顺序访问时每次访问到的都是不同组, 可提高缓存命中率。

7. 因为页<sup>内</sup>偏移在地址翻译中保持不变, 如果使组索引+块内偏移位数=页内偏移, 缓存就可以不等TLB完成地址翻译, 直接用索引位索引, 从而使地址翻译和缓存索引并行, 减少命中时间。

8. 1)  $T = 1.97\% + 110.3\% = 4.27$  个周期

2) 完全随机访问, 可得平均缓存缺失率为  $1 - \frac{64KB}{1GB} = 1 - \frac{1}{2^{14}}$  个  
故平均延迟时  $T = 1.97\% + 110 \cdot (1 - \frac{1}{2^{14}}) \approx 110$  个周期

3) 若程序使用的数据时间局部性和空间局部性高, <sup>处理器</sup>其访问性能将大大提升。由于缓存空间远小于主存, 而访问速度远高于主存, 故局部性高的数据访问可最大化地体现缓存对处理器性能的优化。

4) 设最小命中率为  $x$ ,  $1 \cdot x + 110 \cdot (1 - x) = 105 \Rightarrow x = 4.6\%$

9. 缓存大小(KB)	块大小(B)	相联度	组数量	组索引位数 Bit	标签位数 Bit	偏移位数 Bit	地址位数 32Bit
1	4	2	32	5	21	6	
2	4	8	8	3	23	6	
3	4	全相联	1	0	26	6	
4	16	1	256	8	18	6	
5	16	2	64	6	19	7	
6	64	4	256	8	18	6	
7	64	16	64	6	20	6	
8	64	16	32	5	20	7	



$$10. / 1) 0.22(1-p_1) + 100.22p_1 < 0.52(1-p_2) + 100.52p_2$$

$$\Rightarrow p_1 - p_2 < 0.3\%$$

$$2) 0.22(1-p_1) + k \cdot 0.22p_1 < 0.52(1-p_2) + k \cdot 0.52p_2$$

$$\Rightarrow 0.22p_1 - 0.52p_2 < \frac{0.3}{k-1}$$

11. ①直接映射, 共16个组, 故有4位索引, 题中数据计算得索引分别为1, 5, 1, 5, 5, 5, 5, 5, 故发生5次块替换。

②2路组相联, 8组, 3位索引, ~~标签索引分别为~~ 每组2路, 故共3次块替换

③4路组相联, 4组, 每组4路, 索引均为1, 故共3次块替换

④8路组相联, 2组, 每组8路, 索引均为1, 故共0次块替换

12. 缓存A, 8组, 每组2路, 每个int为4字节, ~~缓存内共有64个int~~, 每个块存放4个int, ~~该array占24个块~~. 而每次循环中根据LRU策略均发生缓存缺失, 故

$$\text{缺失率} = \frac{1}{4} = 25\%$$

缓存B: 16组. 同样每4次循环发生一次缓存缺失, 缺失率25%

13. 
$$\text{int } i; \text{ for } (i=0; i<128; i++) \{$$

$$\text{for } (\text{int } j=0; j<64; j++) \{$$

$$A[i][j] = A[i][j] + 1;$$

$$\}$$

14. 优化前缓存中可存放128个块, 每个块可存放8个int, 循环按第0, 8, 16...个块进行读写。

A中共1024块, ~~除初始128/8=16个块~~, 以8的余数相同的索引为一组, 每组除初始128/8=16

次读写, 每次读写均发生块替换, 故缺失率 
$$\frac{128 \times 8 - 16}{128 \times 8} = \frac{63}{64}$$
, 缺失次数  $64 \times 128 = 8192$





优化后按连续内存读写, 除前128个块, 之后的块均须替换, 故缺失次数为1024次

2) 优化前, 除前128次<sup>读写</sup>~~替换~~, 之后均发生块替换, 即缺失次数为  $128 \times 64 = 8192$

优化后, 与直接映射相同, 缺失次数1024

3) 优化前, 为避免块内数据在还未操作前就被替换, 须缓存有1024块, 即大小为32KB.

优化后, 不存在上述问题, 缓存只要为32B的整数倍即可, 但缓存过小缺失率也会提高.

15.	input				output			
	列0	列1	列2	列3	列0	列1	列2	列3
行0	miss	miss	hit	miss	miss	miss	miss	miss
行1	miss	hit	miss	hit	miss	miss	miss	miss
行2	miss	miss	hit	miss	miss	miss	miss	miss
行3	miss	hit	miss	hit	miss	miss	miss	miss

16. 1) 每个块存放4个int, input的每一行有32个块, 读写顺序为第0, 32, 1, 33...个块.

缓存共有  $\frac{512}{16} = 32$  块, 分为16组, 每组2路. 根据LRU替换策略, 缓存缺失次数为64次, 故命中率为  $\frac{128 \times 2 - 64}{128 \times 2} = \frac{3}{4} = 75\%$

2) 不可以, 按照input的读写顺序, 每4次循环至少发生2次~~块替换~~缓存缺失, 在缓存初始为空情况下命中率不会改善.

3) 可以, 增加块中大小提高了一次读入数据至缓存的数据量, 可以延长发生一次缓存缺失所需的循环次数, 从而提高命中率.

