

5. 1) 需要  $\frac{2^{64} \times 8}{4 \times 2^{20}} = 2^{53}$  字节用于存储页表, 即 512TB

2) 需要  $\frac{2^{64} \times 8}{2 \times 2^{20}} = 2^{37}$  字节用于存储页表, 即 128GB

3) 多级页表通过将整个虚拟地址空间分成多个较小的区域, 从而降低了实际页表存储开销。另外, 多级页表还允许操作系统对于物理内存的分配和回收的分配和回收进行更细粒度的控制, 使得内存管理更加灵活和高效。

⑥ 因为这样可以更好地利用局部性原理。将地址的中间位作为组索引可以将相邻的内存块映射到同一个缓存组中, 从而提高了缓存命中率。同时使用高位作为标签可以避免不同块之间的冲突。

⑦ ① 可以减少虚拟页表内存之间的转换次数, 从而提高访问速度

② 由于程序通常会访问相邻的内存地址, 因此这样可以让缓存更好地利用空间局部性

③ 由于缓存块与页的大小相同, 可以减少额外的硬件设计。

⑧ 存储系统平均访问时间  $t = 1 + 10 \times 0.03 + 105 \times 0.97 = 106.15$  周期

$$10 \times 0.03 + 1 \times 0.97 = 4.27 \text{ 周期}$$

$$2) \frac{2^{20}}{2^{20}} \times 1 + (1 - \frac{2^{20}}{2^{20}}) \times 10 = 110 \text{ 周期}$$

3) 访问同一位置或相邻位置都易命中, 因此利用好局部性可以极大提高处理器的访存性能

4) 设平均缓存命中率为  $y$ , 则  $y + (1-y) \cdot 110 < 105$ ,

$$\text{解得 } y > 4.59\%$$

	Bit	KB	Byte					
9.	1	32	4	64	2	32	3	21 6
	2	32	4	64	8	8	3	23 6
	3	32	4	64	全相联	1	0	25 6
	4	32	16	64	1	256	8	18 6
	5	32	16	128	2	64	6	19 7
	6	32	64	64	4	256	8	18 6
	7	32	64	64	16	64	6	20 6
	8	32	64	128	16	32	5	20 7

10. 1)  $0.22 + 100p_1 < 0.52 + 100p_2$   
 可得条件为  $p_1 - 13 < 0.08$   
 2)  $0.22 + 0.22kp_1 < 0.52 + 0.52kp_2$   
 可得  $0.22p_1 - 0.52p_2 < \frac{0.3}{k}$

```
13. for (int j=0; j<128; ++j) {
    for (int i=0; i<64; ++i) {
        A[j][i]++
    }
}
```

11. ① 直接映射: 替换 5 次  
 ② 2 路组相联: 替换 3 次  
 ③ 4 路组相联: 替换 3 次  
 ④ 8 路组相联: 替换 0 次

14. 优化前的命中率为 0,  
 缺失  $264 \times 128 = 2^{14}$  次  
 优化后缺失率为  $\frac{\frac{1}{3} + 1}{2 \times 10} = 193.75$   
 $= 6.25$

12. A 策略: 缺失率为  $\frac{2^4}{10} = 25\%$   
 B 策略: 缺失率为  $1 - \frac{2}{3} \times 75\% - \frac{1}{3} \times (\frac{99 \times 1}{100} + \frac{100.75}{100}) = 16.75\%$

故优化后缺失  $2^{14} \times 6.25\%$   
 由于是全相联映射,  
 故优化后缺失率为 0  
 (3) 优化前:  $8 \times 128 \times 32B = 32KB$   
 优化后: 32B

15. miss miss hit miss miss miss miss miss  
 miss hit miss hit miss miss miss miss  
 miss hit miss miss miss miss miss miss  
 hit miss hit miss miss miss miss miss

16. 1) 缓存大小为 512 字节, 块大小为 16 字节, 每个块存放 4 个整型数.  
 则第一次访问每个块时会发生强制缺失, 命中率为 75%.  
 2) 不能, 因为块的大小没有改变, 仍然只能存放 4 个数.  
 3) 可以. 因为这样发生强制缺失的次数减少.