

17. 假设一个使用虚拟内存和 L1 缓存的存储系统具有以下特征：

- 内存系统按字节寻址，访存请求每次仅传递一个字节给处理器。
- 虚拟地址长度 14 比特，物理地址长度 12 比特。
- 页大小 64 字节，使用单级页表。
- TLB 拥有 16 个条目，四路组相联。
- L1 缓存物理寻址，块大小 4 字节，共 16 个组，直接映射。

现在 CPU 发起了一次对虚拟地址 0x05a4 的单字节内存加载请求，回答以下问题。

- 若请求发起时，TLB 的部分内容如下表所示。则 TLB 是否发生命中？如果命中，此次内存访问的物理地址是多少？

组号	标签	物理页号	有效位	标签	物理页号	有效位
0	0x0B	—	0	0x1F	—	0
	0x07	0x0D	1	0x02	0x2F	1
1	0x01	0x05	1	0x05	0x0D	1
	0x14	—	0	0x2A	0x16	1
2	0x03	—	0	0x05	0x1C	1
	0x0B	0x07	1	0x00	0x1B	1
3	0x26	0x34	1	0x02	—	0
	0x19	0x2F	1	0x38	—	0

- 该系统的页表有多少个条目？
- 如果 TLB 命中，则使用 1) 得到的物理地址，否则使用物理地址 0x1e4。如果 L1 缓存的内容给如下表所示，则此次访存请求是否命中缓存？如果命中，访存结果是多少？

组号	标签	有效位	块偏移			
			0x0	0x1	0x2	0x3
0	0x1F	0	—	—	—	—
1	0x05	1	0x02	0x09	0xCB	0xA3
2	0x1C	1	0x09	0x55	0x01	0x08
3	0x0D	0	—	—	—	—
4	0x1B	1	0x9B	0xEE	0xE2	0x86
5	0x2F	1	0x00	0x00	0x01	0x00
6	0x07	0	—	—	—	—
7	0x05	1	0x6F	0x23	0xAB	0xD0
8	0x16	0	—	—	—	—

9	0x1C	1	0x63	0x2F	0x1B	0x00
10	0x1C	1	0x28	0x34	0x01	0xC4
11	0x16	1	0x29	0xC8	0x56	0x99
12	0x34	0	—	—	—	—
13	0x34	0	—	—	—	—
14	0x0D	0	—	—	—	—
15	0x07	1	0xE8	0x59	0x04	0x45

(1) 命中，物理地址 0xDA4

(2) 页表项数 = $\frac{2^{14}}{2^6} = 256$

(3) 0xDA4 = $\begin{matrix} \uparrow & \uparrow & \uparrow \\ \text{tag} & \text{index} & \text{offset} \\ 0x36 & & \end{matrix}$

第9组，标签不对，缓存不命中。

18. 一段程序循环往复地按顺序访问 A、B、C、D 四个地址上的数据。考虑一个拥有 2 条目的全相联缓存，回答以下问题。

1) 使用 LRU 替换策略时，填写下表。当程序长时间运行时，缓存的命中率为多少？

访存地址	A	B	C	D	A	B	C	D
way 0	—	A	A	C	C	C	B	B
way 1	—	—	B	B	D	A	A	C
命中?	N	N	N	N	N	N	N	N

2) 提出一种缓存替换策略，使得上述程序可以在该缓存中拥有最大的命中率，并计算该命中率。

(1) 缓存命中率为 $\frac{A}{N} \times \frac{C}{N} \times \frac{C}{N} \times \frac{B}{N} \times \frac{B}{N}$

(2) 不会，认为 FIFO 也为 0%。随机替换可能会提高命中率。

19. 一些处理器引入了“微标签”(microtag)的技术来降低组相联缓存标签匹配过程的时序压力。该技术将地址的标签部分进一步拆分为高位标签 (HTag) 和低位标签 (LTag)，在判断缓存命中与否时，控制器仅取出低位标签进行比较，将匹配的缓存块预测为一次命中并把数据前馈给处理器。在随后的剩余周期内，高位标签被取出并进一步用于判断该预测最终是否构成真正的命中。回答以下问题：

- 1) 低位标签在同一缓存组内通常被要求是唯一的，试说明原因。
- 2) 基于对 1) 的讨论，简要说明该技术的引入对于通常的缓存替换策略有什么影响。
- 3) 考虑到虚拟页偏移和物理页偏移是一致的，为了提高访存性能，系统可以进一步要求地址的低位标签和组索引位完全位于页偏移字段内，这样低位标签的匹配过程就完全不需要经过地址翻译而可以直接进行，后续的高位标签则使用页表翻译后的结果判断是否构成真实命中。基于上述过程，对于 16KB 页大小的内存系统，一个 8KB 大小的四路组相联缓存至多可以拥有几比特的低位标签？

(1) 低位标签的唯一性要求确保不同数据块不会被映射到同一缓存组中，可以减少冲突。

(2) 数据访问更快，减少替换冲突，可以提高命中率，提供更好性能。

(3) 页为 16KB \Rightarrow 有 14 位地址

$$8KB / 1B \times 4 = 2K$$

$$\text{最多可有 } \log_2(2K) = 11 \text{ 位}$$

20. 监听一致性和目录一致性各有什么优缺点？简述缓存一致性的实现代价体现在哪些方面？

监听一致性：

优：实现简单直接，适用于规模较小的系统。

缺：对总线带宽要求高，可扩展性低。

目录一致性：

优：可扩展性好，适用于大规模多处理器系统。

缺：需要更多硬件支持来维护和管理目录。

实现代价：

总线带宽和延迟、调度和冲突处理。