

# 作业 4

SA20225085 朱志儒

1. 假设磁盘块大小为 8 KB，块中存储 200 字节的定长记录，块首部只包括一个 8 字节的模式指针和一个偏移量表。对于插入块内的每条记录，在偏移量表中都增加一个 2 字节的指针指向该记录。假设每天向块内插入 4 条记录（空间不足时允许插入部分记录后结束全部操作），删除 2 条记录。假设每天的删除记录操作总是发生在插入记录之前，删除记录使用一个“删除标记”代替记录在偏移量表中的指针。给定一个磁盘块，如果刚开始块是空的，则几天后不能再向该块内插入记录？此时，该块内一共有多少条记录？

解：第一天磁盘块新增数据大小为：

$$8 + 4 \times 2 + 4 \times 200 = 816 \text{ B}$$

之后，磁盘块每天插入数据的大小为：

$$2 \times 2 + 2 \times 200 = 404 \text{ B}$$

第 19 天磁盘存储的数据大小为：

$$816 + 18 \times 404 = 8088 \text{ B}$$

而磁盘的总容量为：

$$8 \times 1024 = 8192 \text{ B}$$

显然，20 天后不能再向该块内插入记录。

此时，该块内的记录数为：

$$4 + 2 \times 18 + 1 = 41$$

2. 假设我们采用 LRU 作为缓冲区置换策略，当我们向 Buffer Manager 发出一个读页请求时，请讨论一下：

（1）如果页不在缓冲区中，我们需要从磁盘中读入该页。请问如何才能在缓冲区不满的时候快速地返回一个 free 的 frame？请给出至少两种策略，并分析一下各自的时间复杂度。

（2）如何才能快速地判断所请求的页是否在缓冲区中？如果请求的页在缓冲区中，如何快速返回该页对应的 frame 地址？请给出至少两种策略，并分析一下各自的时间复杂度。

解：（1）策略一：

使用位图记录 frame 的使用情况，1 表示 frame 被占用，0 表示 frame 为 free。遍历位图找出一个标记为 0 的 frame，返回这个 free 的 frame，时间复杂度为 $O(n)$ 。

策略二：

维护一个空闲链表，表中记录空闲的 frame，每当请求来临时，返回表中第一个 free 的 frame，时间复杂度为 $O(1)$ 。

(2) 策略一：

采用 BCB 数组存储缓冲区的使用情况，遍历一遍该数组，查找所请求的页是否在缓冲区，若在则返回该页对应的 frame 地址。时间复杂度为 $O(n)$ 。

策略二：

采用 Hash 表维护缓冲区的使用情况，使用拉链法解决冲突，Hash 函数：

$$H(k) = \text{pageid} \% \text{buffersize}$$

根据 page\_id 计算 Hash 值，然后在 Hash 表中查找对应的 BCB，若找到则返回 BCB 中存储的 frame\_id。时间复杂度为 $O\left(\frac{n}{k}\right)$ ，其中 $k = \text{buffersize}$ 。