

20337025-崔璨明-12

学号	姓名	章号
20337025	崔璨明	12

12.3

关系 r_1 需要 $20000/25=800$ 个块，关系 r_2 需要 $45000/30=1500$ 个块。设内存的页数量为 m ，若 $m>800$ ，则可以使用简单的嵌套循环连接即可，需要 $800+1500$ 次块传输，2次磁盘搜索；因此考虑 $m\leq 800$ 页的情况：

a.

若使用 r_1 作为外层关系，需要 $20000*1500 + 800 = 30000800$ 次块传输， $20000+800=20800$ 次磁盘搜索。

若使用 r_2 作为外层关系，需要 $45000*800 + 1500 = 36001500$ 次块传输， $45000+1500=46500$ 次磁盘搜索。

b.

若使用 r_1 作为外层关系，需要的块传输次数为： $\lceil \frac{20000/25}{M-1} \rceil * (45000/30) + (20000/25)$ ，需要磁盘搜索次数 $\lceil 2 * \frac{800}{M-1} \rceil$

若使用 r_2 作为外层关系，需要的块传输次数： $\lceil \frac{45000/30}{M-1} \rceil * (20000/25) + (45000/30)$ ，需要磁盘搜索次数 $\lceil 2 * \frac{1500}{M-1} \rceil$

c.

设 r_1 和 r_2 最初并没有在连接键上排序，则排序的总成本为：

$$Bs = 1500(2\lceil \log_{M-1}(\frac{1500}{M}) \rceil + 2) + 800(2\lceil \log_{M-1}(\frac{800}{M}) \rceil + 2)$$

总的块传输次数： $Bs + 1500 + 800$

需要磁盘搜索次数：

$$2\lceil \frac{800}{M} \rceil + \lceil \frac{800}{b_b} \rceil * (2\lceil \log_{\lceil \frac{M}{b_b} \rceil - 1} \frac{800}{M} \rceil) + 800 + 2\lceil \frac{1500}{M} \rceil \\ + \lceil \frac{1500}{b_b} \rceil * (2\lceil \log_{\lceil \frac{M}{b_b} \rceil - 1} \frac{1500}{M} \rceil) + 1500 + 2\lceil \frac{1200}{b_b} \rceil$$

b_b 为分配给归并段或者每个关系的缓冲块。

d.

r_1 较小，因此用 r_1 作为构造用输入， r_2 作为探查用输入。

若 $M > 800/M$ ，则不需要递归分区，则块传输次数为 $3 * (1500 + 800) = 6900$ ，磁盘搜索次数为 $2(\lceil \frac{800}{b_b} \rceil + \lceil \frac{1500}{b_b} \rceil)$

若 $M \leq 800/M$ ，则块传输次数为 $2(1500 + 800)\lceil \log_{M-1}(800) - 1 \rceil + 1500 + 800$ ，磁盘搜索次数为 $2(\lceil \frac{800}{b_b} \rceil + \lceil \frac{1500}{b_b} \rceil)(\lceil \log_{M-1} 800 \rceil - 1)$