1. 1) 块嵌套循环连接

∵ 有10页缓冲区

∴ 可以每次读入1页Course，9页student

$\lceil \frac{500}{9} \rceil = 56$，　　$20 \times 56 = 1120$　　$1120 + 20 = 1140$

∴ 需要读出 1140 页

2) 索引嵌套循环连接

假设有100条Course记录满足

索引总读页数 $\lceil \log_2(100) \rceil + 9 = 17$

∵ 每条满足条件的course都要重新扫描一次student

$500 \times 100 = 50000$

∴ 读 50017页，写 0页

∴ 总 I/O 为 50017页

3) 排 排序归并连接

a. 排序：

对于Student，每次读10页，排序后写回

∴ 读写分别 $\frac{500}{10} = 50$页，共100页

对于Course，可直接一次性排序，I/O = 0

∴ 排序I/O为40页

b. 归并

对于Student：需合并 $\lceil \frac{50}{9} \rceil = 6$次

$5 \times 90 + 50 = 500$页

∴ 总 I/O 为 2000页

对于Course，归并需20页

∴ 归并I/O为 2020页

∴ 块嵌套循环连接 I/O最低

1)

2. $\Pi_{title, gpa}(\sigma_{credits<4}(Course) \bowtie Student)$

1) $\Pi_{title, gpa}((\Pi_{title, dept}(\sigma_{credits<4}(Course))) \bowtie (\Pi_{dept, gpa}(Student)))$

3) $\Pi_{title, gpa}((\Pi_{dept, gpa}(Student)) \bowtie (\Pi_{title, credits}(\sigma_{credits<4}(Course))))$

3. 1) $P \to S \to R$ ~~$P \to S \to R$~~

   ~~$S \to P \to R$~~ $S \to R \to P$

   $R \to S \to P$ ~~$R \to P \to S$~~

2) ① $P \to S \to R$ $600 + 600 \times 400 + (600 + 600 \times 400) \times 1 = 480600$

   ② $S \to P \to R$ 同上，约为 $400 + 24000 \times 2 = 480400$

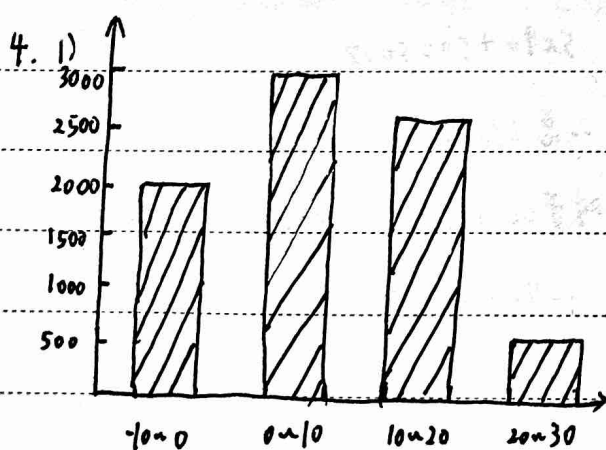   ③ $R \to S \to P$ $7 + 7 \times 400 + 2800 \times 600 = 168\overset{2}{\cancel{3}}807$

   ~~④ $P \to R \to S$ $600 + 600 \times 50 + 600 \times 50 \times 400 = 12030600$~~

   ⑤ $S \to R \to P$ $3200 + 2800 \times 600 = \cancel{240800}$ 1683200

   ~~⑥ $R \to P \to S$ $7 + 7 \times 600 + 4200 \times 400 = 1684207$~~

∴ 3) $S \to R \to P$ 最优，先连接较小的表，减少了内表的记录

4. 1)

2) ① $1500 + 1250 = 2750$ ② 假设均匀分布 $500 \times \frac{1}{10} = 50$

优：

3) 分区划分容易，不需要提前知道数据分布   缺：差值太大会有误差

   宽度一致，便于计算   对于点查询很难估计

可能误差：$10 \sim 20$ 和 $20 \sim 30$ 之间差值过大 i 点查询效果差