



Atenção: A entrega será via e-mail (brayner@dc.ufc.br). Como o trabalho pode ser feito por até três alun*s, o nome do arquivo enviado deve ter o formato: SGBD_2921_1AP_<nome_participa1>_<nome_participa2>_<nome_participa3>

- 1) Considere a seguinte variação para o operador físico *block nested-loop join* para implementar o operador de junção da álgebra relacional:

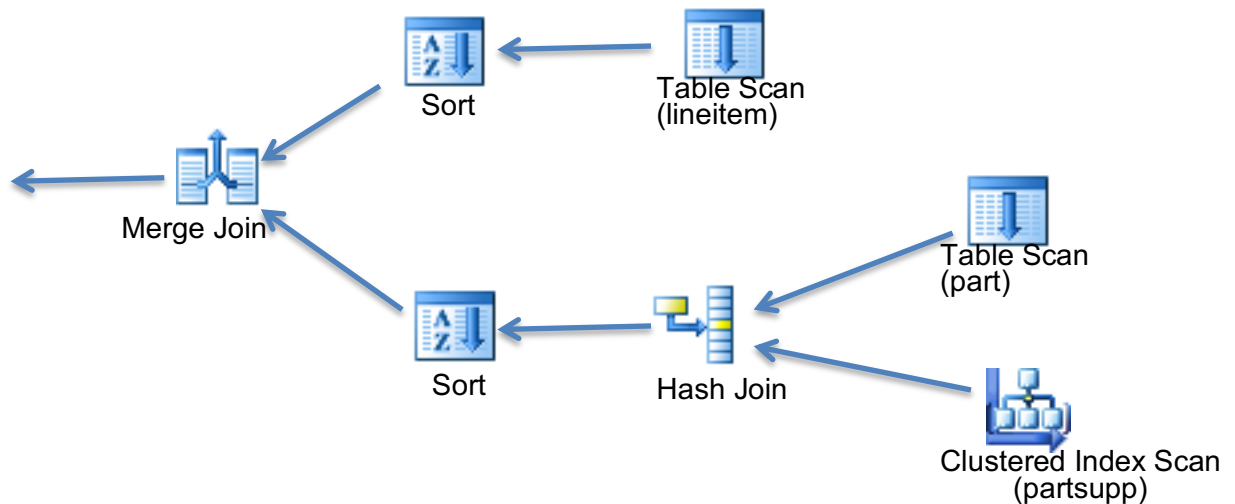
```
P=1
Faça enquanto não-fim de s
  i=1
  Ler página P de s
  Faça enquanto i<=k e não-fim de s
    Ler próxima página de s
    i=i+1
  Fim;
  P=P+k
  Para cada página  $P_r$  da tabela r faça
    Para cada página  $P_s$  da tabela s em buffer faça
      Para cada tupla  $t_s$  em  $P_s$  faça
        Para cada tupla  $t_r$  em  $P_r$  faça
          Se o par  $(t_r, t_s)$  satisfaz a condição de junção
            inclua  $t_r.t_s$  ao resultado
          Fim-Se;
        Fim;
      Fim;
    Fim;
  Fim;
```

Fim:
Calcule a estimativa de custo para o operador definido acima. Para tanto considere que r e s possuem P_r e P_s páginas, respectivamente, onde $P_r > P_s$. Justifique sua resposta.

- 2) Considere a seguinte consulta submetida no banco de dados TPC-H (fator de escala 1):

```
Select l_partkey, l_orderkey, p_partkey, p_name, ps_partkey, l_quantity
from lineitem inner join partsupp inner join part on ps_partkey=p_partkey
on l_partkey=ps_partkey
where l_quantity>20
order by l_partkey
```

As tabelas lineitem e part não possuem índices. A tabela partsupp tem um índice primário, com chave de busca ps_suppkey. O plano gerado pelo processador de consultas do SQL Server foi o seguinte:



Ao analisar o plano, você identificou que poderia reduzir o tempo de execução da consulta. Para tanto, certas ações deveriam ser executadas sobre o banco de dados TPC-h. Indique que ações são estas e apresente o novo plano gerado.

- 3) A estratégia de *merge-join* requer que as tabelas envolvidas na operação de junção estejam ordenadas pelo atributo de junção. Uma variante desta estratégia, o ***merge-join híbrido***, requer que apenas uma das tabelas esteja ordenada pelo atributo de junção. Contudo, deve existir um índice secundário para a outra tabela envolvida na junção, definido sobre o atributo da condição de junção e implementado como uma árvore B^+ . Dessa forma, pode-se implementar o operador físico *merge-join híbrido*, lendo-se a tabela ordenada e executando um *index scan* sobre a estrutura de índice da outra tabela. Escreva um pseudo-código para a implementação deste operador físico e Determine a estimativa de custo para que o processador de consultas execute a operação algébrica $R \bowtie_{R.C=S.C} S$, através do operador físico *merge-join híbrido*. As tabelas R e S apresentam os seguintes esquemas: $R(A, B, C)$ e $S(C, D, E)$. Considere ainda que foram definidos um índice primário em S sobre o atributo $S.C$ e um índice secundário em R sobre o atributo $R.C$. A árvore do índice secundário de R é de ordem n e apresenta uma altura h . Suponha que não existem níveis de indireção para o índice secundário de R . $FS(C,R)$ e $FS(C,S)$ representam o fator de seletividade do atributo C nas tabelas R e S , respectivamente.

- 4) Considere a definição do índice *ind_emp_01* no SQL Server, apresentada a seguir:
Create clustered index ind_lineitem on lineitem (l_partkey) with (pad_index=on, fillfactor=1)

Determine a altura da árvore B^+ construída para implementar *ind_lineitem*, considerando que a tabela *lineitem* tem cardinalidade igual a $6,002 \times 10^6$ tuplas. As páginas têm capacidade de 8Kbytes, onde 8000 bytes representam o espaço utilizável para armazenar dados. A chave de busca *l_partkey*, do tipo *int*, tem tamanho de 4 bytes. Cada ponteiro da árvore necessita de 4 bytes para ser armazenado e que cada tupla da tabela *lineitem* tem tamanho de 137 bytes. O parâmetro *fillfactor* (juntamente com *pad_index*) representa a taxa de



Universidade Federal do Ceará
Centro de Ciências
Departamento de Computação
Sistemas Gerenciadores de Banco de Dados – CK0117 (T02)
Lista para composição da nota da 1ª. AP
Prof. Dr.-Ing. Angelo Brayner

Prazo para Entrega: Até às 23.59 de 30/07/2021

preenchimento (quantidade de entradas, em %) de um nó da árvore, durante a criação do índice. O parâmetro *clustered* determina que o índice a ser criado é do tipo primário.