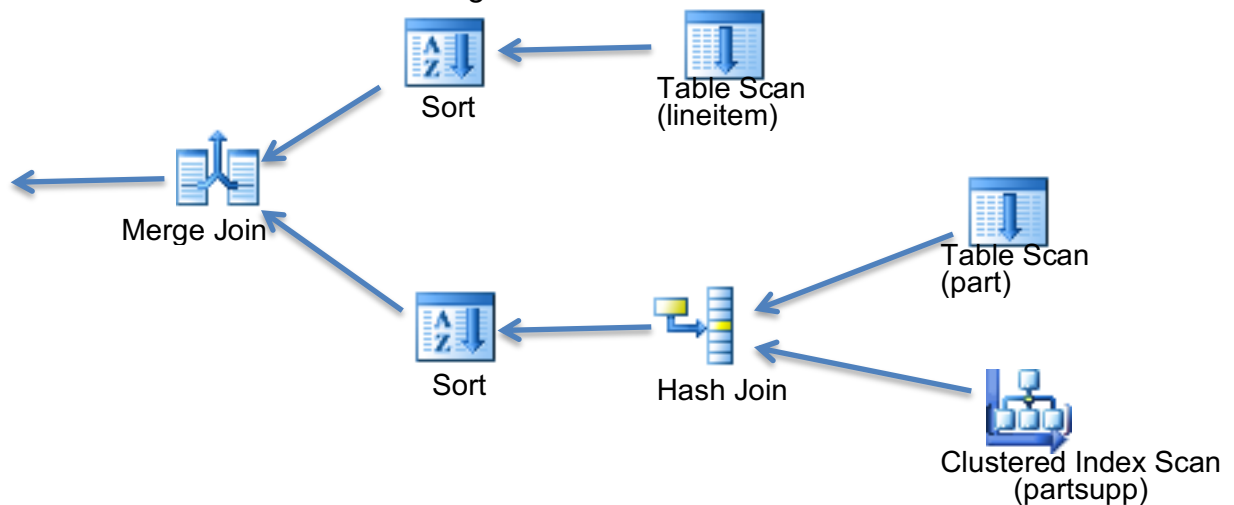




- 1) Considere a seguinte consulta submetida no banco de dados TPC-H (fator de escala 1):

```
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.

- 2) Explique como a ordem de uma árvore B+ **está relacionado com o** tamanho da página do banco de dados e porque isso influencia diretamente o desempenho de operações de leitura e escrita em disco.
- 3) Analise o desempenho da árvore B+ em dois cenários distintos:
- Consulta por intervalo de valores contínuos
  - Consultas frequentes por chaves aleatórias
- Compare a performance e explique por que a árvore B+ é preferida em ambos os casos em bancos de dados com armazenamento em disco.

- 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 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.

- 5) Determine a estimativa de custo para calcular uma operação  $R \bowtie S$ , considerando que foi definido um índice hash sobre o atributo de junção da tabela R. Considere que não há *bucket overflow*.