UNIVERSIDADE ESTADUAL DE SANTA CRUZ (UESC)

Criada pela Lei 6.344, de 05.12.1991,
e reorganizada pela Lei 6.898, de 18.08.1995 e
pela Lei 7.176, de 10.09.1997

CET091 – Banco de Dados II

Prof. Dr. Marcelo Ossamu Honda

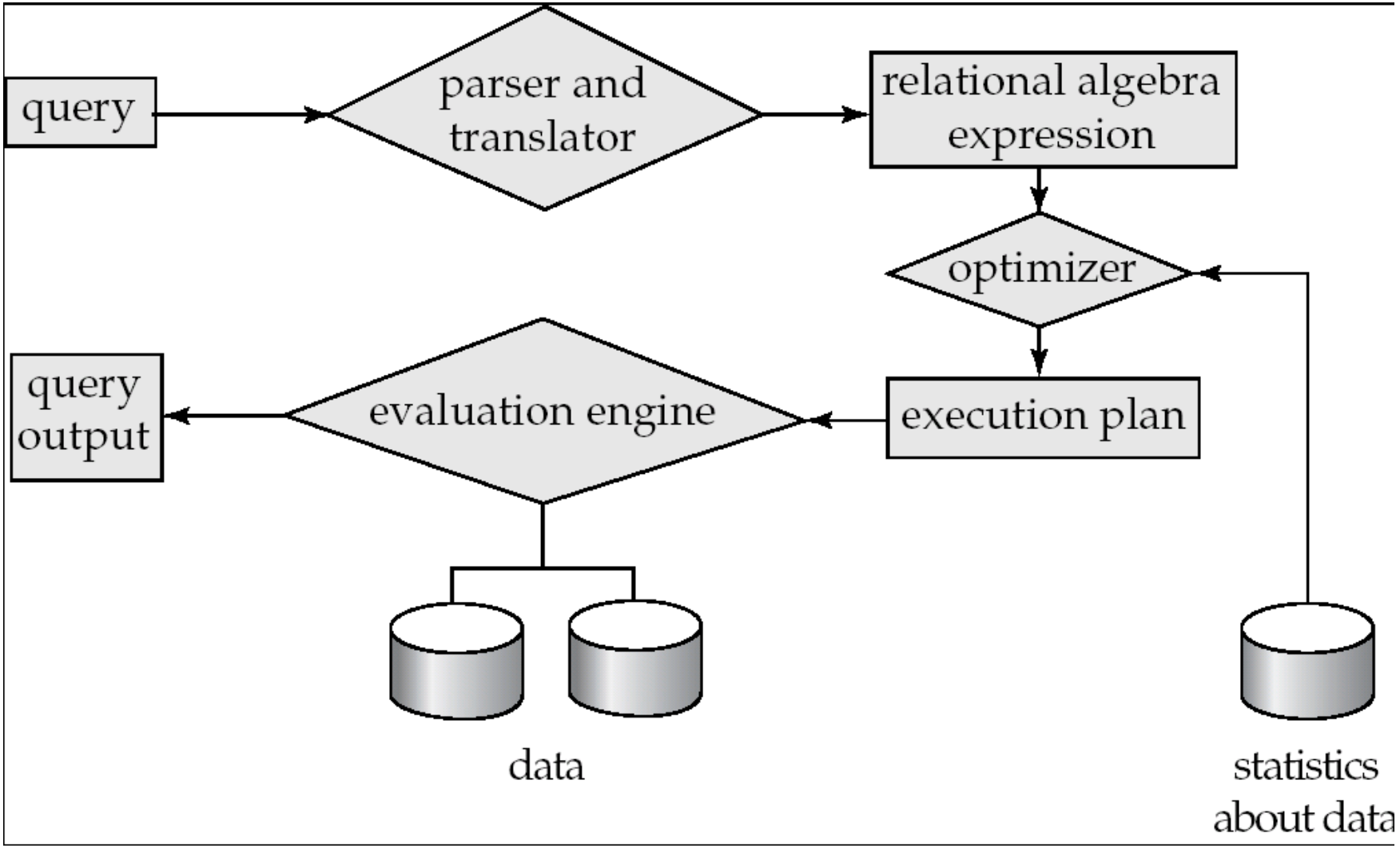
Departamento de Ciências Exatas e Tecnológicas (DCET)
mohonda(at)nbcgib(.)uesc(.)br

Processamento da Consulta

Introdução

- Processamento da consulta:
  - Refere-se ao conjunto e atividades envolvidas na extração de dados de um banco de dados;
    - Tradução de consultas em linguagens de banco de dados de alto nível para expressões que podem ser usadas no nível físico do sistema de arquivos;
      - Análise e tradução;
    - Série de transformações de otimização de consulta;
      - Otimização;
    - Avaliação real das consultas;
      - Avaliação;

Processamento da Consulta



Visão Geral

- Etapa de Análise e tradução:
  - Tradução das consultas em um representação interna;
    - Álgebra relacional;
  - Analisador:
    - Verificar a sintaxe;
    - Nomes dos atributos;
    - Verifica as relações;

Visão Geral

- Etapa de Avaliação:
  - O mecanismo de execução de consulta apanha um plano de avaliação de consulta, executa esse plano e retorna os resultados das consultas;
    - Select saldo from conta where saldo < 2500
      - $\sigma_{Saldo < 2500}{}^{(\P\ saldo(conta))}$ ou ¶ saldo ($\sigma_{Saldo < 2500}{}^{(conta)}$)
  - Plano de avaliação (ou execução) da consulta;
    - Primitivas de avaliação;
      - Uma operação da álgebra relacional anotada com instruções sobre como avaliá-la;
        - ¶ saldo
        - $\sigma_{Saldo < 2500}$; usar índice 1
        - conta

Visão Geral

- Etapa de Otimização:
  - Determinar o custo de cada operação;
    - Difícil de ser determinado;
    - Calculado uma estimativa;
  - Otimização de consulta;
    - O sistema de banco de dados deve ser capaz de construir um plano de avaliação de consulta que minimize o custo da avaliação da consulta;
      - Os diferentes planos de avaliação para determinada consulta podem ter diferentes custos;

Medidas de Custo da Consulta

Medidas de Custo da Consulta

- Tempo de resposta para um plano de avaliação de consulta, ou seja, a quantidade de tempo necessária para executar o plano;
- Recursos diferentes:
  - Acesso a disco;
    - Determinante em sistemas grandes;
    - Diferentes valores para leitura e escrita;
  - Tempo de CPU;
    - Difícil de estimar (detalhes de baixo nível);
  - Memória;
    - Geralmente considerado o pior dos casos:
      - Poucos buffers em memória;
      - Sempre o buffer precisa ser lido do disco;
  - Custo de comunicação;
    - Sistemas distribuídos ou paralelos;

Operação de Seleção

Operação de Seleção

- Varredura do arquivo;
  - Considerado o operador de menor nível para acessar dados;
  - São algoritmos de busca que localizam e apanham registros que cumprem uma condição de seleção;

Operação de Seleção

- Algoritmos Básicos:
  - A1 - Busca Linear;
    - O sistema varre cada bloco do arquivo e testa todos os registros para verificar se satisfazem a condição de seleção;
      - Necessário uma busca inicial, para determinar o primeiro bloco do arquivo;
        - Buscas extras quando o arquivo não é contíguo;
      - Atributo chave, a busca pode terminar sem examinar todos os registros;
    - Pode ser aplicado a qualquer arquivo;
      - Independente da classificação do arquivo;
      - Disponibilidade de índices;
      - Tipo da operação de seleção;
    - Mais lento dos algoritmos;

Operação de Seleção

- Algoritmos Básicos:
  - A2 - Busca Binária;
    - Seleção usando atributo chave;
      - O sistema realiza a busca binária sobre os blocos do arquivo;
      - Cada um desses acessos de bloco exige uma busca de disco, além de uma transferência de bloco e o custo de busca e latência do disco;
    - Seleção usando atributo não-chave;
      - Custo da busca usando atributo chave;
      - Custo da leitura de blocos extras;

Operação de Seleção

- Seleções usando índices:
  - Algorítimos de busca que usam índices são considerados varreduras de índices;
  - Os índices oferecem:
    - Acesso rápido, direto e ordenado;
      - Índice Primário (índice agrupado), o arquivo físico mantem os registros ordenados (próximo);
      - Índice Secundário, não é um índice primário;
  - Os índices impõem uma sobrecarga de acesso aos blocos;

Operação de Seleção

- Algoritmos de busca que usam um índice:
  - A3 (índice primário, igualdade sobre chave):
    - Exemplo árvore B+:
      - Altura da árvore (I/O) + I/O para ler o registro;
  - A4 (índice primário, igualdade sobre atributo não chave):
    - Exemplo árvore B+:
      - Altura da árvore (I/O) + I/O para ler o(s) registro(s) que contenham os registros;
  - A5 (índice secundário, igualdade):
    - Sobre chave, similar ao A3;
    - Não chave, cada registro pode residir em um bloco diferente, o que pode resultar em uma operação de I/O por registro;
      - Cada operação de I/O exigindo uma busca e uma transferência de bloco;
    - Resultado pode ser pior do que uma busca linear;

Operação de Seleção

- Seleções envolvendo comparações:
  - A6 (índice primário, comparação):
    - Exemplo árvore B+ para comparação A > v ou A ≥ v;
      - Pesquisado o valor (v) no índice para encontrar a primeira tupla no arquivo;
      - Varredura no arquivo até o final, custo pode ser determinado como o caso A4;
    - Exemplo árvore B+ para comparação A < v ou A ≤ v;
      - Varredura simples, até o valor (v) da tupla ser encontrada;
      - Não é necessário uma pesquisa por índice;
  - A7 (índice secundário, comparação):
    - Para operadores: >, ≥, < e ≤ v;
      - Os índices secundários utilizam ponteiros aos registros reais;
        - Pode exigir operações extras de I/O, pois registros consecutivos podem estar em diferentes blocos de disco;
        - Em buscas com um grande número de registros pode se tornar mais dispendioso do que o uso de uma busca linear;

Operação de Seleção

- Implementação de seleções complexas:
  - Seleções conjuntivas;
    - SELECT * FROM EMPREGADOS
    - WHERE SALARIO > 5000 AND SEX='F' AND DEP=100;
  - A8 (Seleção conjuntiva utilizando um índice):
    - Determinar se um caminho de acesso está disponível para um atributo em uma das condições simples;
      - Se sim, utilizar um dos métodos de operação de seleção para selecionar os registros que satisfazem a condição;
    - Testar no buffer, se cada registro satisfaz as condições simples restantes ou não;
    - O custo final depende do método de operação escolhido;

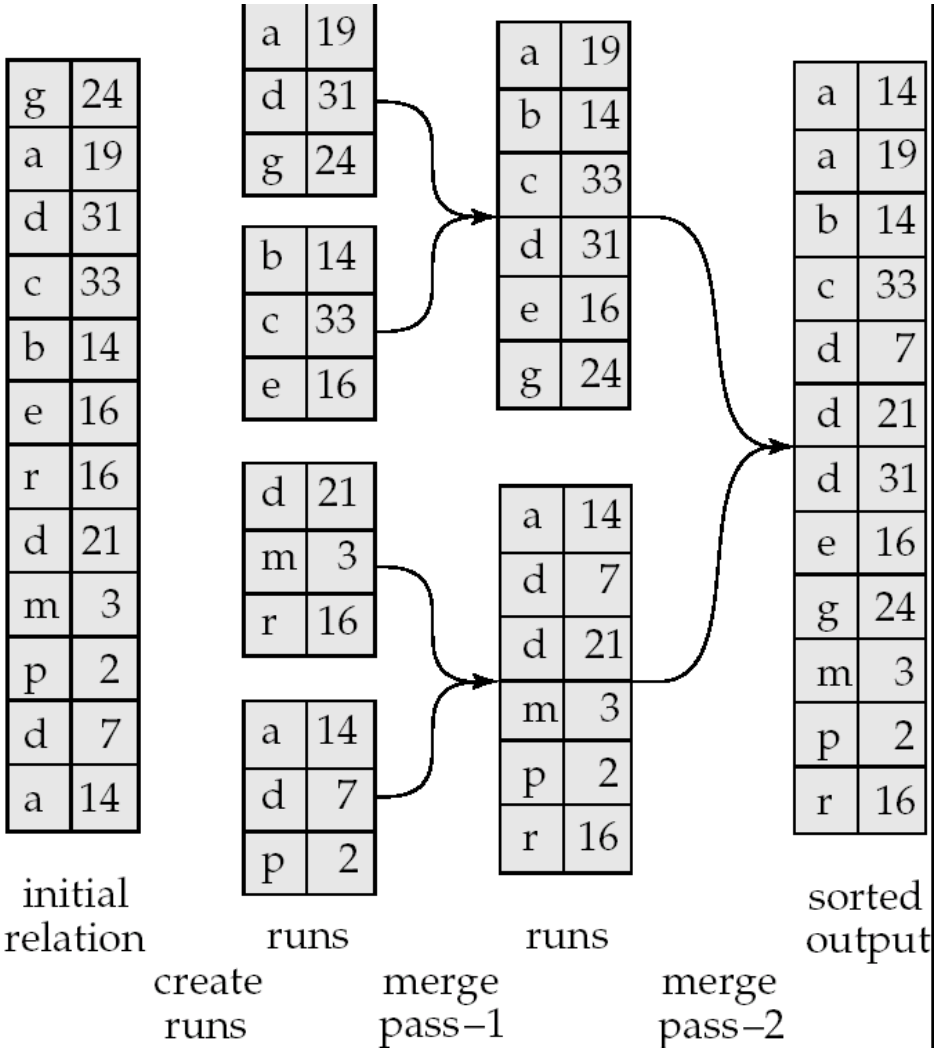
Classificação

Classificação

- Papel importante nos sistemas de banco de dados:
  - Consultas classificadas;
  - Processar operações relacionais eficientemente;
    - Exemplo: junções;
- Classificação de relações:
  - Cabem em memória;
    - Técnicas de classificação padrão;
  - Relações que não cabem em memória;
    - Classificação externa;

Classificação

- Algoritmo sort-merge externo;
  - Merge de N vias;



Operação de Junção

Operação de Junção

- Junção Natural:
  - Operação binária que permite combinar certas seleções e um produto cartesiano em uma única operação;
- A equi-junção;
  - Operação de junção horizontal entre duas tabelas usando uma comparação por igualdade entre a(s) coluna(s) comum(ns);

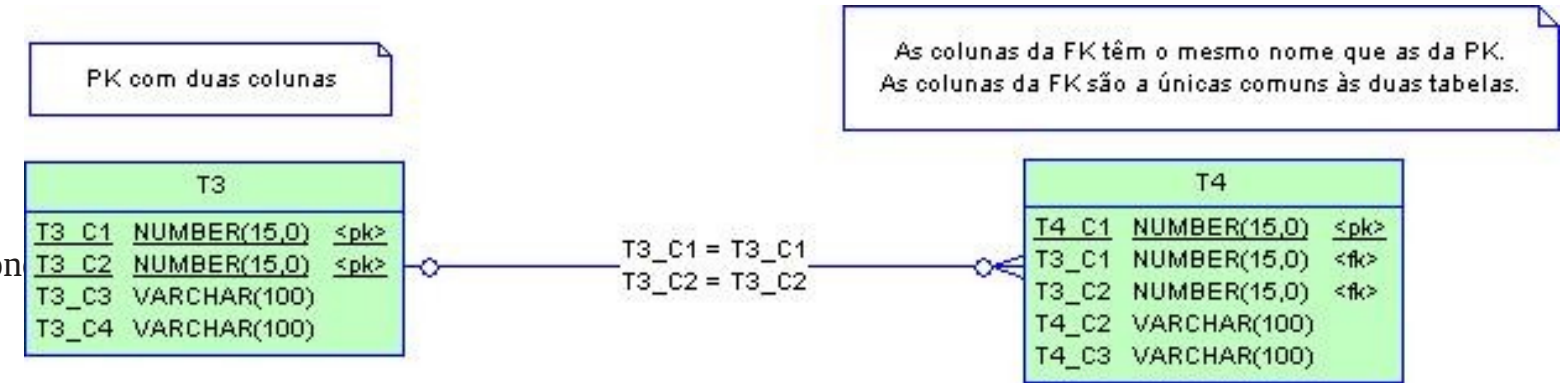
Operação de Junção

- Junção Natural:

```
SELECT *
FROM t3
NATURAL JOIN t4;
```

- A equi-junção;

```
SELECT *
FROM t3, t4
WHERE t3.t3_c1=t4.t3_c1
  AND t3.t3_c2=t4.t3_c2;
```



Operação de Junção

- Operação de junção:
  - Algoritmos para calcular a junção de relações e seus respectivos custos;
    - Depositante ►◄ Cliente
  - Junção de loop:
    - Aninhado:

```
for each tuple t_r in r do begin
  for each tuple t_s  in s do begin
        test pair (t_r,t_s) to see if they satisfy the join condition θ
        if they do, add t_r • t_s to the result.
  end
end
```

      - Basicamente consiste em um par de loops (for) aninhados:
        - Relação externa;
          - Delimita a relação interna;
        - Relação interna;
          - Quando uma relação cabe inteiramente em memória principal, é recomendável utilizar-la como relação interna;
      - Não exige índices;
      - Independente da condição de junção;
      - Junção natural;
        - Eliminação de atributos repetidos por projeção;
      - Custoso pois examina cada para de tuplas nas duas relações;

Operação de Junção

- Operação de junção:
  - Junção de loop (cont.):
    - Aninhado em bloco:
      - Processa as relações com base em cada bloco;
        - No lugar de cada tupla;
    - Aninhado indexado:
      - Utiliza de índices existentes ou índices temporários;
        - O custo é melhor nas relações onde o índice é existente;

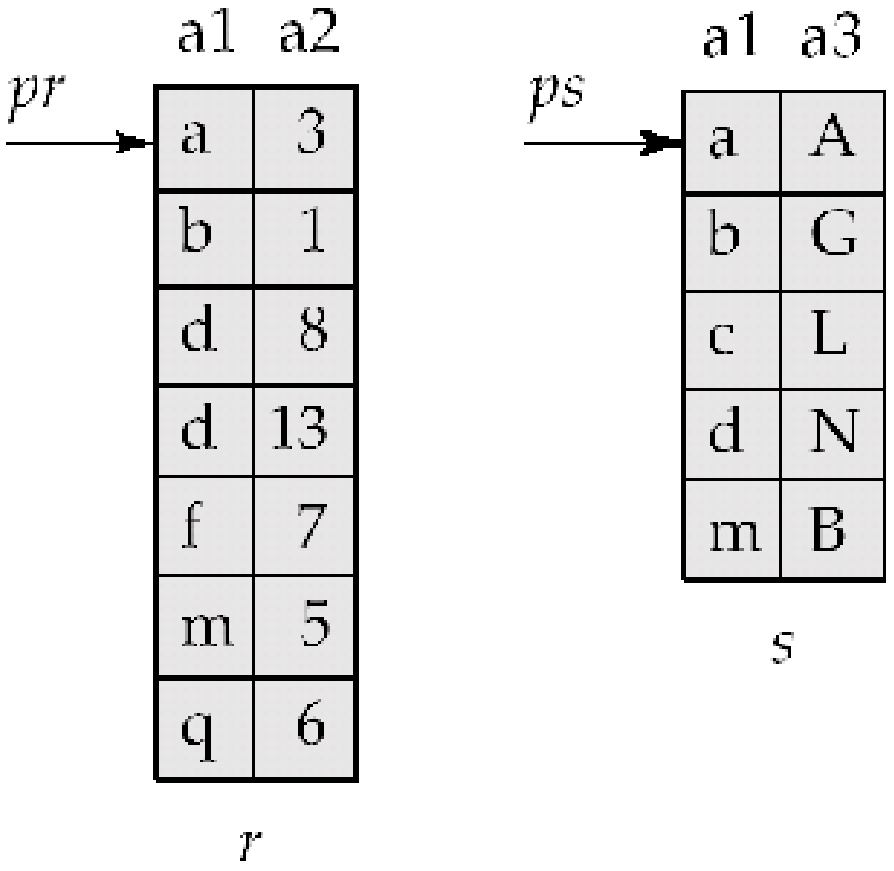
```
for each block B_r of r do begin
    for each block B_s of s do begin
        for each tuple t_r in B_r do begin
            for each tuple t_s in B_s do begin
                Check if (t_r,t_s) satisfy the join condition
                if they do, add t_r • t_s to the result.
            end
        end
    end
end
```

Operação de Junção

- Operação de junção:
  - Melhora o desempenho de junções de loop Aninhado e Aninhado em bloco:
    - Atributos da união formarem uma chave sobre a relação interna;
    - Usar o maior tamanho que pode caber em memória;
      - Espaço para buffer da relação interna e de resultado;
    - Varrer o loop interno alternadamente;
      - Reutilizar os dados em buffer;
    - Índice disponível no atributo de junção do loop interno;
      - Utilizar métodos mais eficientes;

Operação de Junção

- Operação de junção:
  - Junção merge (sort-merge):
    - Junções naturais e equijunções:
    - Atributo de junção deve ter índex em ambas relações;
    - Pode utilizar processo semelhante ao algoritmo sorte-merge;
    - Um dos conjuntos de tuplas, deve caber em memória principal;
    - Como é preciso apenas uma leitura em ambos arquivos, o método é eficiente;

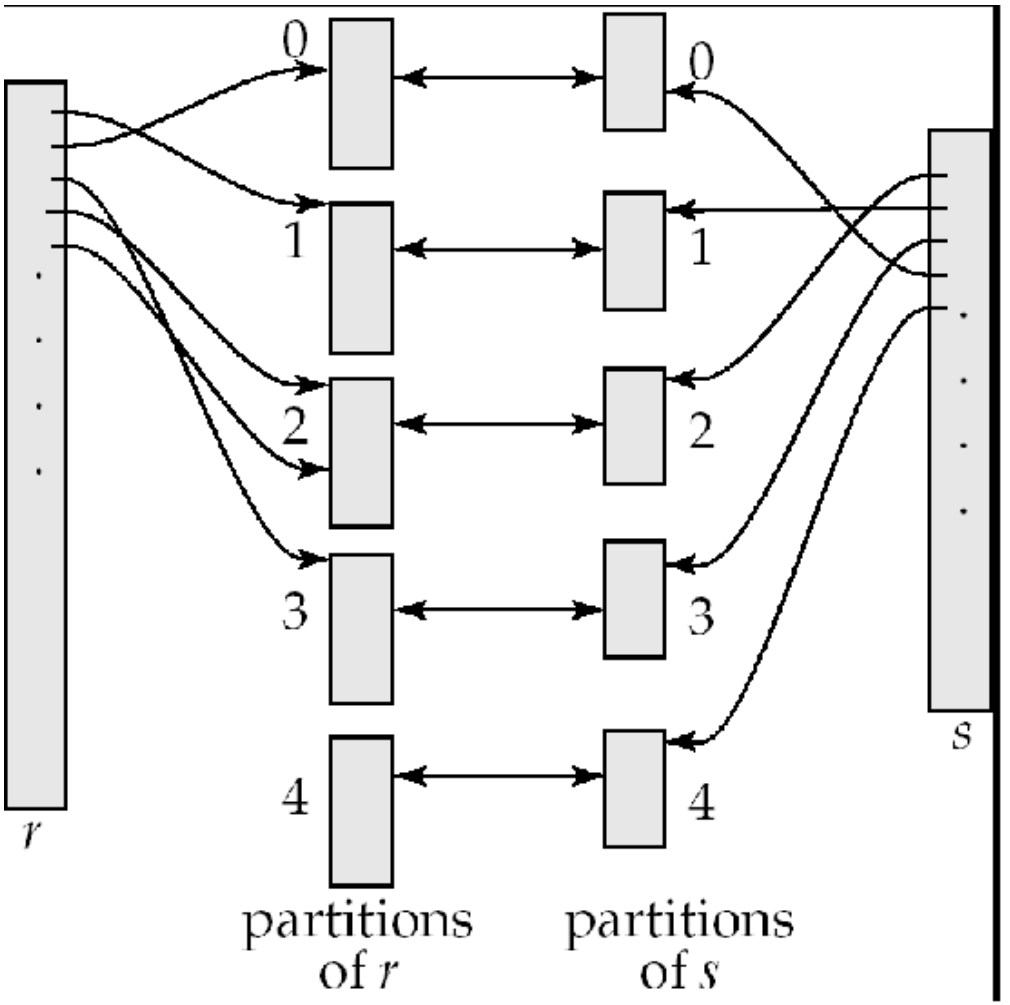


Operação de Junção

- Operação de junção:
  - Junção merge (sort-merge) hibrido:
    - Quando uma das relações é classificada;
      - Mas a outra relação possui um índice de árvore B+ secundário sobre os atributos de junção;
    - Mescla a relação classificada com as entradas de folha do índice de árvore B+;
      - Tuplas da relação classificada e endereços para as tuplas da relação não classificada;
      - Então é feito a classificação sobre os endereços das tuplas da relação não classificada;

Operação de Junção

- Operação de junção:
  - Junção hash:
    - Junções naturais e equijunções:
    - Uma função hash é utilizada para particionar as tuplas em ambas relações;
      - A ideia básica é particionar as tuplas de cada uma das relações em conjuntos que tem o mesmo valor de hash nos atributos de junção;



Avaliação de Expressões

Avaliação de Expressões

- É a avaliação de uma expressão que contenha várias operações;
  - A maneira mais simples seria avaliar uma operação de cada vez, na ordem apropriada;
- Métodos:
  - Materialização:
    - O resultado de cada avaliação é materializado em uma relação temporária para uso subsequente;
      - Relações temporárias devem ser gravadas em disco (desvantagem), a menos que sejam pequenas;
  - Canalização:
    - Avaliar várias operações simultaneamente em uma canalização (pipeline);
    - Os resultados de uma operação são passados para a seguinte sem a necessidade de armazenar uma relação temporária;

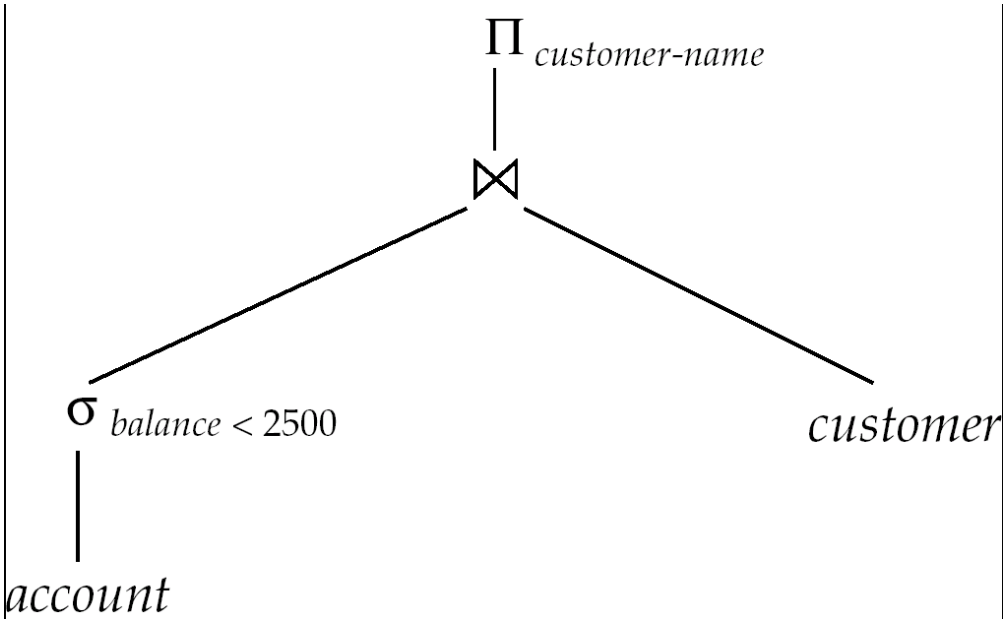
Avaliação de Expressões

- Avaliação Materializada:
  - Examinando a representação gráfica da expressão em uma árvore de operadores;
    - O início se dá pelas operações de nível mais baixo da árvore;
  - Consulta original:

¶ nome_cliente σ Saldo < 2500 (conta) ►◄ cliente

  - Primeira etapa:

σ Saldo < 2500 (conta)



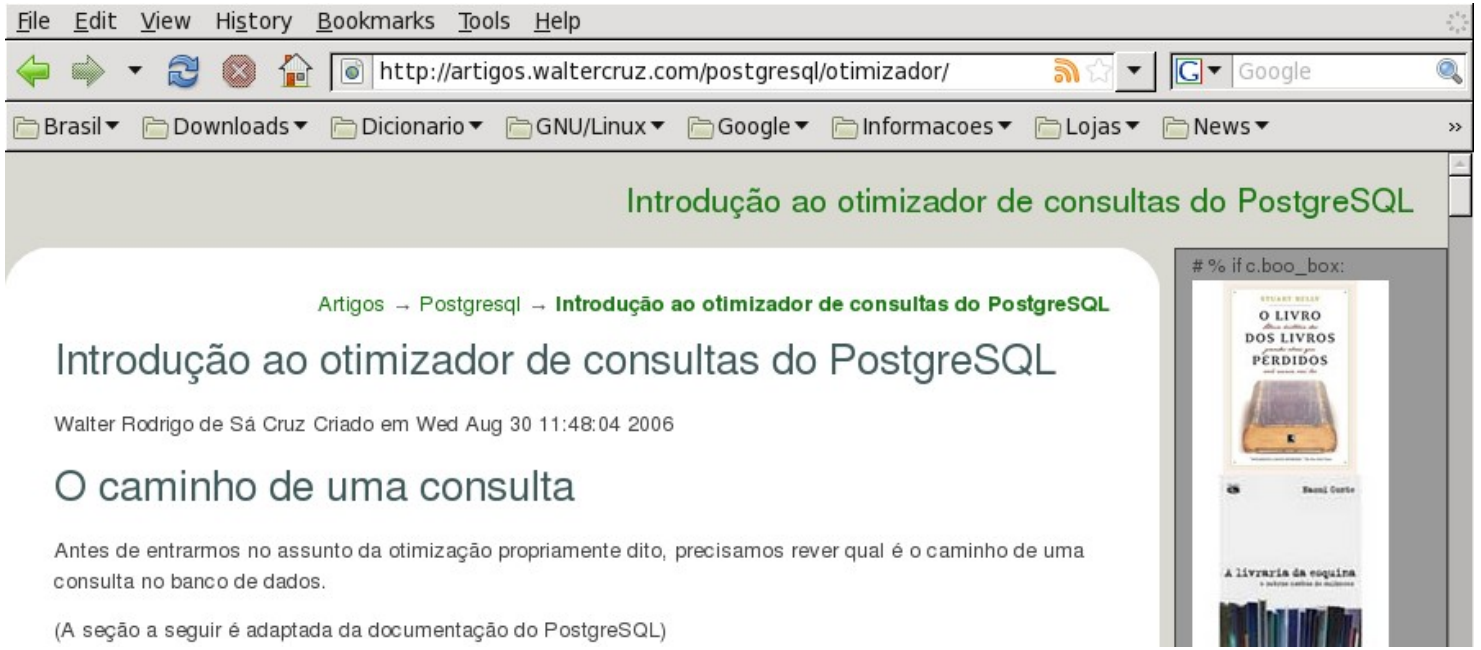
Avaliação de Expressões

- Avaliação Materializada:
  - Para avaliar o custo de uma expressão:
    - Somar os custos de todas as operações;
    - Custo de gravar os resultados intermediários;
  - Buffer duplo:
    - Uso de dois buffers, com um continuando a execução do algoritmo enquanto o outro está sendo gravado, permite que o algoritmo seja executado mais rapidamente;

Avaliação de Expressões

- Canalização:
  - Evita a escrita dos resultados de muitas subexpressões em disco, usando resultados na expressão pai mesmo quando estão sendo gerados;
  - Avaliação canalizada:
    - Implementação:
      - Controlada por demanda:
        - Aguarda ser requisitado novo conjunto de tuplas;
      - Controlada por produtor:
        - Armazena em um buffer de saída (até conter espaço);
    - Nem sempre é possível aplicar a canalização;
      - Junção Merge;

Laboratório



18.6. Query Planning

18.6.1. Planner Method Configuration

These configuration parameters provide a crude method of influencing the query plans chosen by the query optimizer. If the default plan chosen by the optimizer for a particular query is not optimal, a temporary solution can be found by using one of these configuration parameters to force the optimizer to choose a different plan. Turning one of these settings off permanently is seldom a good idea, however. Better ways to improve the quality of the plans chosen by the optimizer include adjusting the *Planner Cost Constants*, running *ANALYZE* more frequently, increasing the value of the default_statistics_target configuration parameter, and increasing the amount of statistics collected for specific columns using `ALTER TABLE SET STATISTICS`.

Prof. Dr. Marcelo Ossamu

Referências

- Ramez Elmasri e Shamkant B, Navathe, Sistemas de Banco de Dados, Pearson Addison Wesley, 2005;
- Abraham Silverschatz, Henry F. Korth e S. Sudarshan, Sistema de Banco de Dados, Editora Campus, 2006;
- PostgreSQL 8.3.6 Documentation, by The PostgreSQL Global Development Group, Copyright © 1996-2008 The PostgreSQL Global Development Group;
- http://artigos.waltercruz.com/postgresql/otimizador/

Referências

- KORTH, H. F., SILBERSCHATZ, A. Sistema de Banco de Dados, Makron Books
- DATE, C. J. Introdução a Sistemas de Banco de Dados. Tradução da 7ª. Edição Americana. Editora Campus.
- DATE, C.J. Bancos de Dados, Tópicos Avançados, Editora Campus
- GARCIA-MOLINA, H. ULLMAN, J.D., WIDOM, J. Implementação de Sistemas de Bancos de Dados. Editora Campus
- Manuais Técnicos e Livros voltados para Bancos de Dados específicos.