

A decorative graphic on the left side of the slide, consisting of a network of white lines and small circles on a blue gradient background, resembling a circuit board or data flow diagram.

ESTRUTURA DE PROCESSAMENTO

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Consulta – pode ter sua resposta computada por uma variedade de métodos (geralmente)
- Usuário (programador) – sugere uma estratégia para achar a resposta, independentemente de ser a estratégia mais eficiente
- SGBD – responsável por transformar a consulta realizada pelo usuário em uma consulta equivalente mais eficiente

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Característica

- melhoria da estratégia para processamento de uma consulta
- não apresenta uma estratégia ótima
- porém apresenta uma estratégia eficiente

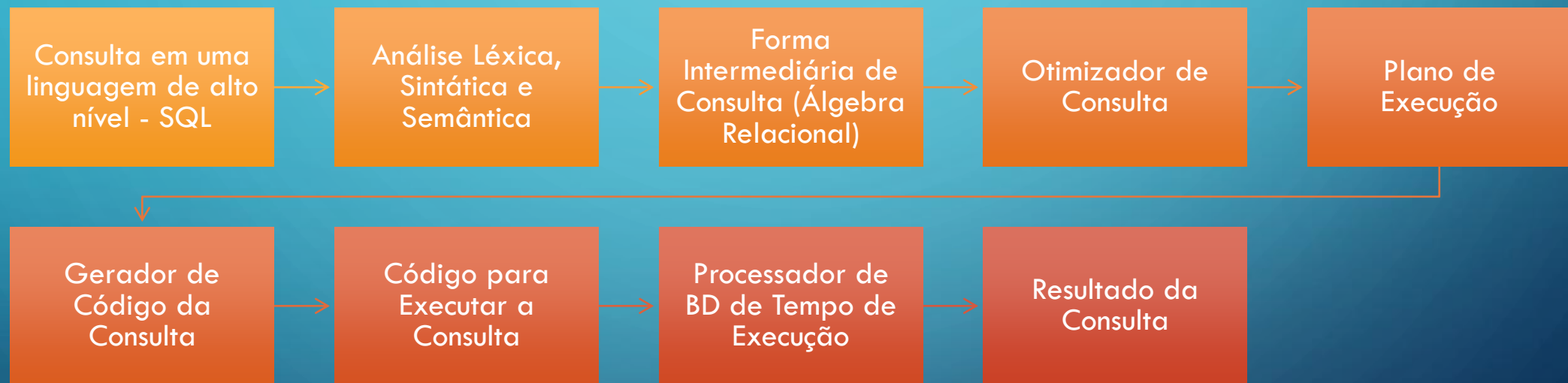
- Modelo relacional

- facilita a otimização da consulta
- permite que uma consulta seja expressa inteiramente em uma linguagem de consulta relacional (i.e., SQL) sem o uso de uma linguagem hospedeira

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Técnica utilizada para processar, otimizar e executar consultas de alto nível
- Objetivo: – produzir uma estratégia de consulta para recuperar o resultado da mesma
- Estratégia de consulta é um plano para:
 - Executar a consulta
 - Acessar os dados
 - Armazenar resultados intermediários

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS



PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Análises (léxica, sintática, semântica)
 - Verificam a sintaxe da consulta
 - Verificam se os nomes das relações da consulta são os mesmos nomes presentes no BD
 - Substituem referências ao nome de uma visão por expressões da álgebra, a fim de computar essa visão
- Forma intermediária de consulta
 - Utiliza estrutura de dados de árvore ou grafo (árvore de consulta ou grafo de consulta)

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Módulo de otimização de consulta
 - Produz um plano de estratégia de execução
 - Indica qual o plano de execução
 - Gera o código da consulta para executar tal plano
- Processador do BD de tempo de execução
 - Executa o código da consulta (compilado ou interpretado)
 - Produz o resultado da consulta

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Fase 0
- Nível de SQL
 - Consiste em transformar a consulta expressa em SQL em uma árvore de consulta expressa em álgebra relacional

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Fase 1
- Nível de álgebra relacional
 - Consiste em aplicar heurísticas para converter uma árvore de consulta em uma árvore de consulta equivalente
 - Consiste em encontrar uma expressão que seja equivalente à expressão dada, mas que seja mais eficiente na sua execução

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Fase 2
- Consiste na seleção de uma estratégia detalhada para o processamento da consulta
 - Como a consulta será executada
 - Quais índices serão escolhidos
 - Qual a ordem de processamento das tuplas

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Árvore de Consulta
- Estrutura de árvore que corresponde a uma expressão da álgebra relacional
 - Representação
 - Nós folhas - relações de entrada para a consulta
 - Nós internos - operações da álgebra relacional
 - Indica uma ordem específica das operações durante a execução de uma consulta

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Execução

- efetuar uma operação do nó interno sempre que os seus operandos estiverem disponíveis
- substituir o nó interno pela relação que resulta da execução da operação
- final da execução:
 - o nó raiz é executado
 - a relação resultado para a consulta é produzida

- O otimizador tem que incluir regras de equivalência entre as expressões da álgebra relacional que podem ser aplicadas à árvore inicial. As regras da otimização da consulta por heurística, utilizam expressões de equivalência para transformar a árvore inicial em final, resultando em uma árvore de consulta otimizada.

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Exemplo - Dado o seguinte comando

```
SELECT LNAME
```

```
FROM EMPLOYEE, WORKS_ON, PROJECT
```

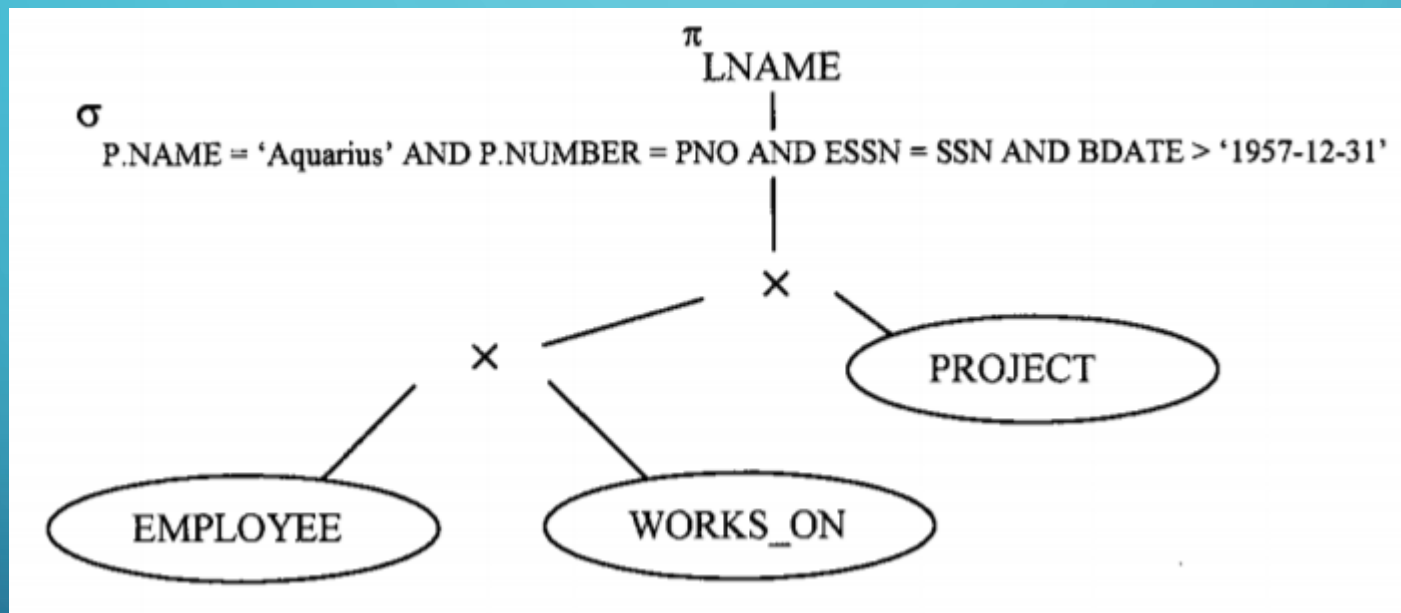
```
WHERE PNAME = 'Aquarius' AND
```

```
PNUMBER = PNO AND ESSN = SSN AND
```

```
BDATE > '1957-12-31'
```

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

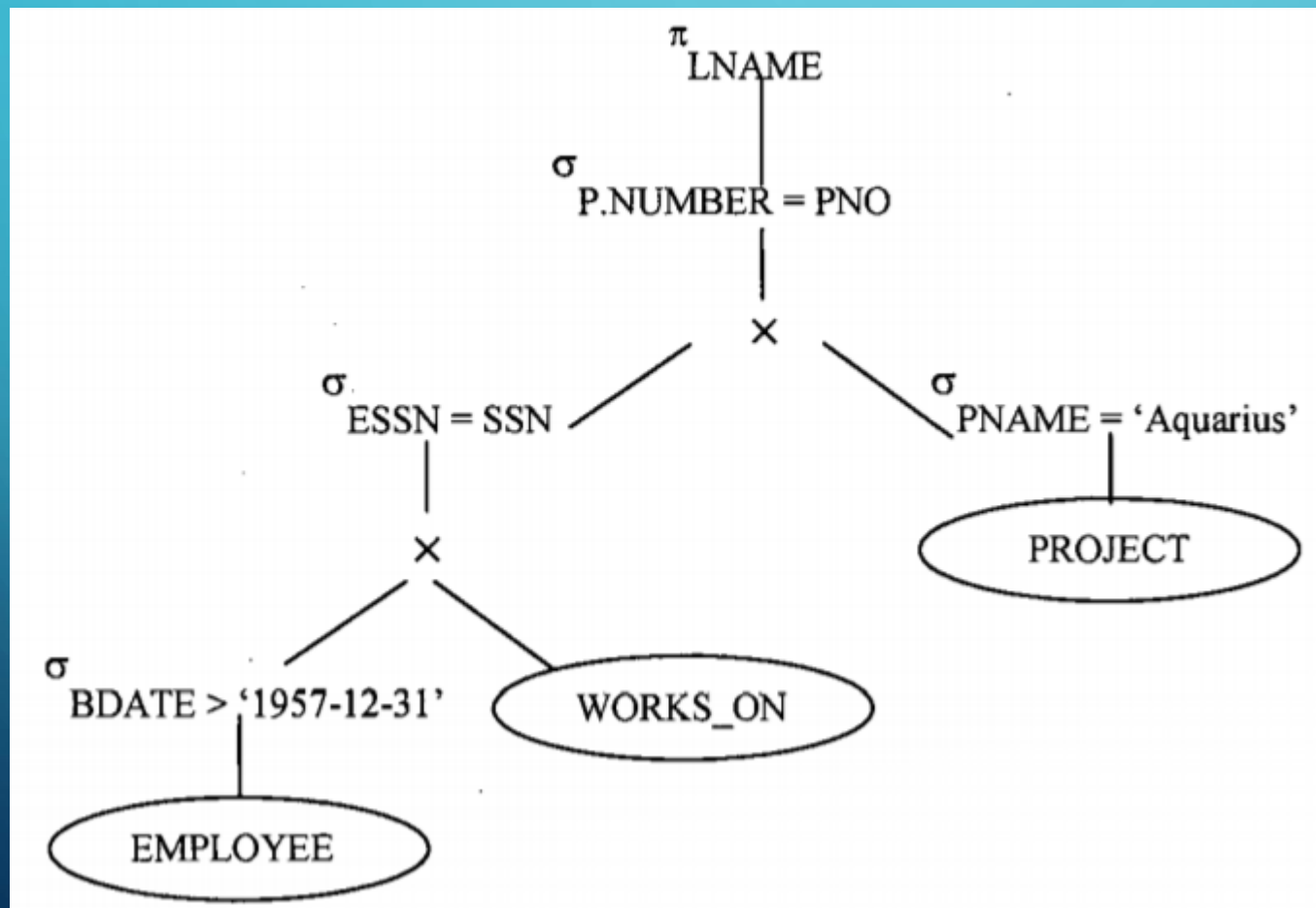
- Árvore de Consulta Inicial



Executando diretamente esta árvore, primeiramente obtêm-se um arquivo bastante grande que contém o produto cartesiano dos arquivos `EMPLOYEE`, `WORKS ON`, e `PROJECT`. Entretanto, esta consulta solicita somente um registro da relação `PROJECT` (projeto Aquarius) e somente um registro da relação `EMPLOYEE` (data de nascimento > 1957-12-31)

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

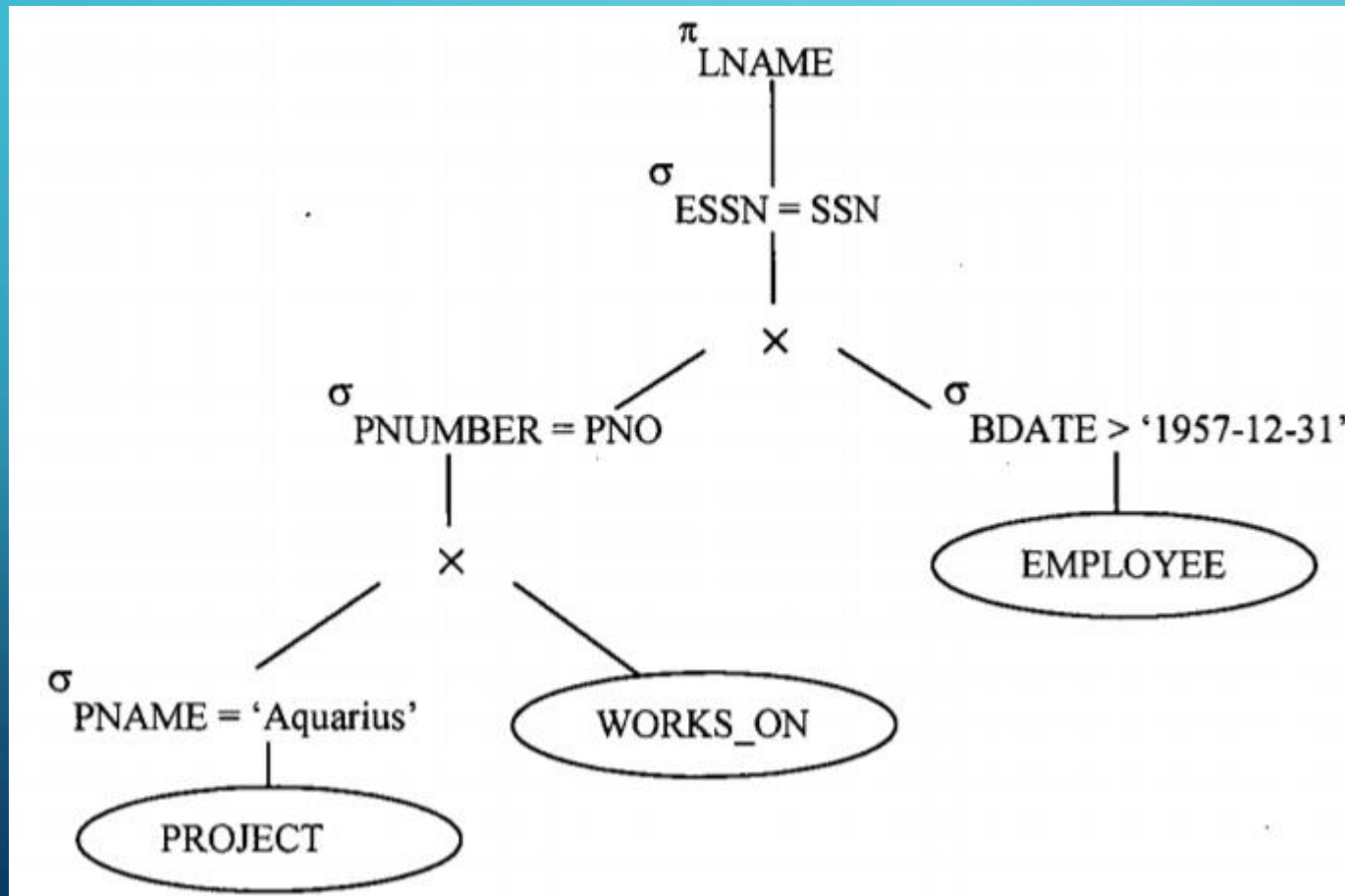
- Árvore de Consulta Melhorada 1



Inicialmente, nesta árvore, aplica-se as operações SELECT para reduzir o número de tuplas que aparecem no produto cartesiano.

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

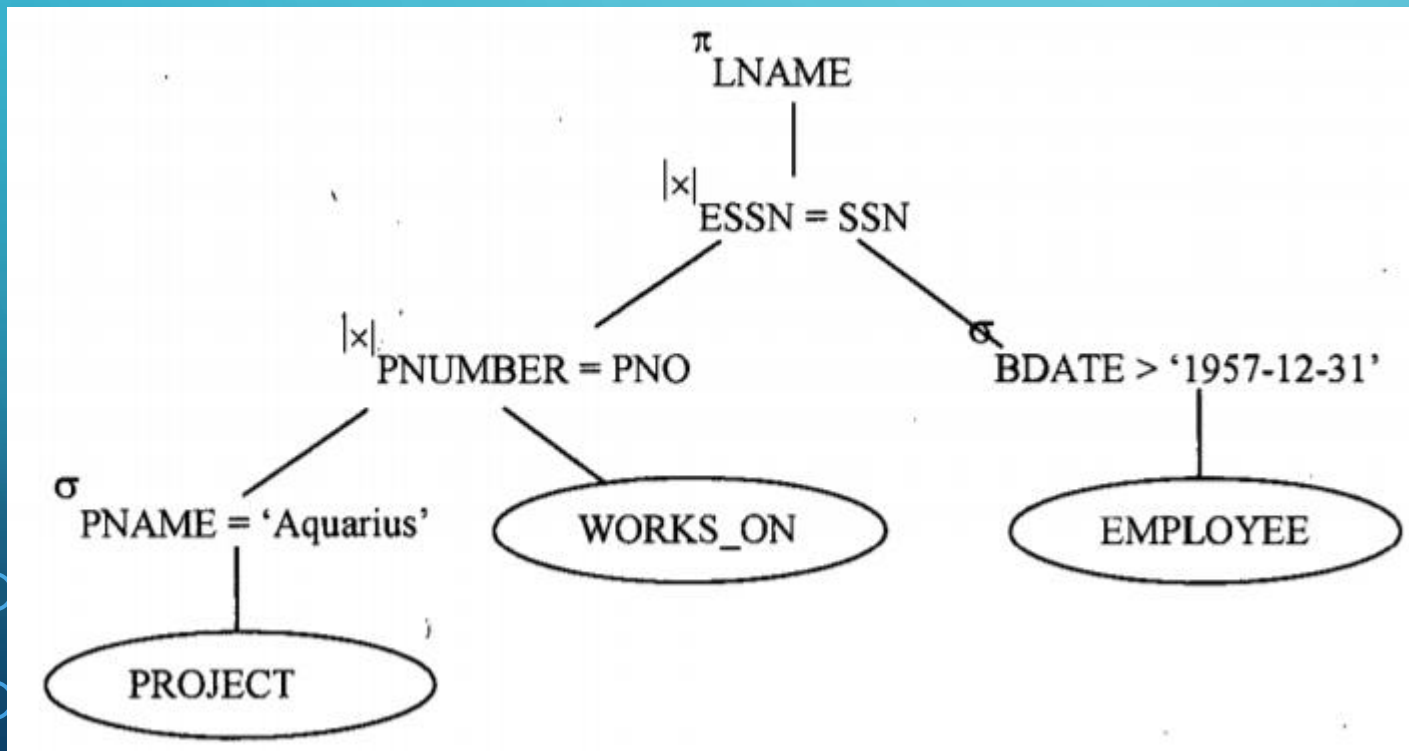
- Árvore de Consulta Melhorada 2



Para que uma melhora adicional possa ser obtida, é necessário trocar na árvore, as posições da relação EMPLOYEE e PROJECT

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

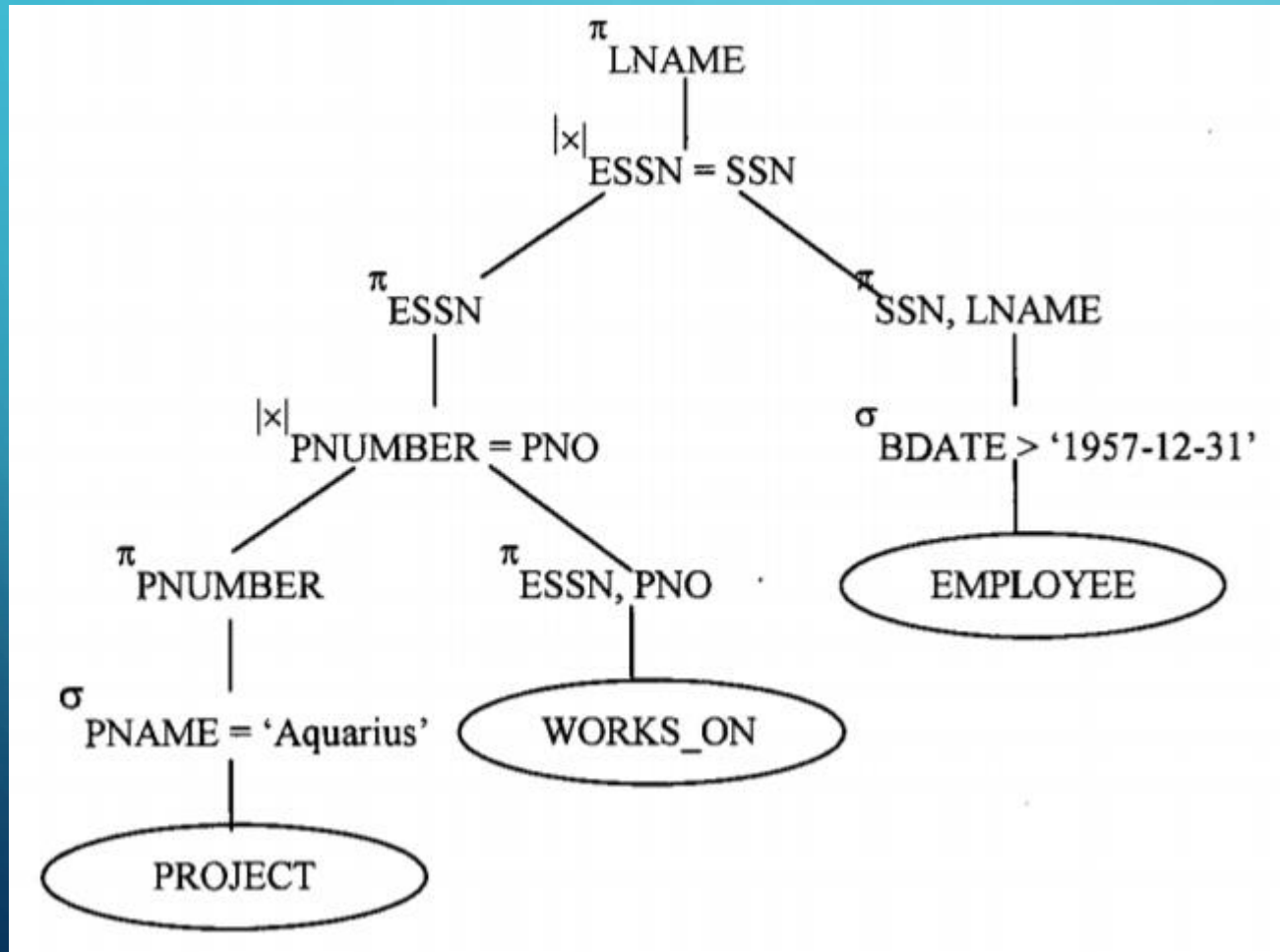
- Árvore de Consulta Melhorada 3



Neste caso, a informação PNUMBER é um atributo chave da relação PROJECT, e consequentemente a operação de seleção da relação PROJECT recobrará um único registro. Mais adiante pode-se melhorar a árvore de consulta, substituindo qualquer operação do produto cartesiano pela operação seguida de uma condição de junção

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Árvore de Consulta Melhorada 4



Dada uma expressão da álgebra relacionai, é função do otimizador de consulta propor um plano de avaliação da consulta que gere o mesmo resultado da expressão fornecida e que seja uma maneira menos onerosa de gerar o resultado (ou que, pelo menos, não seja muito mais cara que a maneira mais barata).

Para encontrar o plano de avaliação de consulta menos caro, o otimizador precisa gerar planos alternativos que produzam o mesmo resultado da expressão dada e escolher o plano menos caro [Silberschatz et al., 1999].

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Heurísticas Básicas

- Aplicar primeiro as operações que reduzem o tamanho dos resultados intermediários
 - operações de seleção - reduzem o número de tuplas
 - operações de projeção - reduzem o número de atributos
- Aplicar primeiro as operações de seleção e de junção mais restritivas
 - reordenar os nós folha da árvore de consulta
 - evitar a operação de produto cartesiano
 - ajustar o restante da árvore de forma apropriada

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Otimização de Consultas baseada no Custo
- Características
 - utiliza técnicas de otimização tradicionais que percorrem o espaço solução para um problema
 - identifica uma solução que minimiza uma determinada função custo
 - depende de estatísticas armazenadas pelo SGBD sobre as relações e sobre os índices

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Componentes da Função Custo

- Custo de acesso à memória secundária – custo para buscar, ler e escrever blocos de dados que residem em disco
- Custo de armazenamento – custo para armazenar quaisquer arquivos intermediários gerados pela estratégia de execução da consulta
- Custo de computação – custo para a realização de operações em memória principal (i.e., buffers)

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Componentes da Função Custo

- Custo do uso da memória – custo relacionado ao número de buffers de memória principal necessários durante a execução da consulta
- Custo de comunicação – custo de transmitir uma consulta e os seus resultados do site do banco de dados até o site ou terminar na qual a consulta foi originada

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Componentes da Função Custo

- Característica
- BD volumosos - minimizar o custo de acesso à memória secundária
- BD pequenos - minimizar o custo de computação
- BD distribuídos - minimizar também o custo de comunicação

PROCESSAMENTO E OTIMIZAÇÃO DE CONSULTAS

- Fonte:
- Banco de Dados – Processamento e Otimização de Consultas, Profa. Dra. Cristina Dutra de Aguiar Ciferri (<http://wiki.icmc.usp.br/images/a/a6/SCC578920131-proconconsultas.pdf>)