

Universidade Federal da Paraíba  
Centro de Ciências e Tecnologia  
Coordenação de Pós-Graduação em Informática

**MAPESGBD**  
**Uma Metodologia de Apoio à Decisão**  
**para Seleção de SGBD**

Adna Magnólia Dantas Silva

DIA  
01.11.65 (045)  
5386m


Campina Grande - PB  
Agosto - 1996



**Adna Magnólia Dantas Silva**

**MAPESGBD - Uma Metodologia de Apoio à Decisão para Seleção de SGBD**

Dissertação apresentada no Curso de  
MESTRADO EM INFORMÁTICA  
da Universidade Federal da Paraíba,  
em cumprimento às exigências para  
obtenção do Grau de Mestre.

**Área de Concentração: Ciência da Computação**

Maria de Fatima Queiroz Vieira Turnell  
**(Orientadora)**

Campina Grande  
Agosto - 1996



## Ficha Catalográfica

Silva, Adna Magnólia Dantas

S586M

MAPESGBD – Uma Metrologia de Apoio à Decisão para Seleção de SGBD  
CCT/COPIN da UFPB, 1996

86p il.

Dissertação (mestrado) – Universidade Federal da Paraíba, Centro de Tecnologia,  
Coordenação de Pós-Graduação em Informática, Campina Grande, 1996. Orientadora:  
Maria de Fátima Queiroz Vieira Turnell.

1. Sistema de Informação e Banco de Dados, 2. Banco de Dados. I. Título

**CDU681.3.07B**

**MAPESGBD - UMA METODOLOGIA DE APOIO À DECISAO PARA  
SELEÇÃO DE SGBD**

**ADNA MAGNÓLIA DANTAS SILVA**

**DISSERTAÇÃO APROVADA EM 09.08.96**

*Maria de Fátima Q. Vieira Turnell*  
**PROF. MARIA DE FÁTIMA Q. V. TURNELL, Ph.D**  
**Presidente**

*Marcus Costa Sampaio*  
**PROF. MARCUS COSTA SAMPAIO, Dr.**  
**Examinador**

*Ulrich Schiel*  
**PROF. ULRICH SCHIEL, Ph.D**  
**Examinador**

*Geovane Cayres Magalhães*  
**PROF. GEOVANE CAYRES MAGALHÃES, Dr.**  
**Examinador**

**CAMPINA GRANDE - PB**

*A minha avó Rita Maria de Lima,  
em memória.*

***“O homem para vencer não deve  
temer coisa alguma, nem  
mesmo uma idéia nova.”***

## *Agradecimentos*

A DEUS, que me abençoou com a faculdade da inteligência para que eu escalasse mais um degrau do conhecimento.

A minha família, especialmente aos meus pais (Ednaldo e Maria), por plantarem a semente da esperança, sempre confiantes no meu potencial.

A minha Orientadora Maria de Fatima Turnell, que como uma verdadeira mestra exerceu sublimemente seu sacerdócio.

A Vladimir, que com sua nobreza de espírito me apoiou nos momentos de cansaço.

A meus amigos (Bel, Cristina, Eliane, Goreth, Iana, Josimar, Kissia, Marcia, Nevinha, Ramiro, Stanley, Roberta, Verônica), especialmente a Claudinha e Flávia, por terem me compreendido e ajudado nos momentos difíceis dessa jornada.

Aos professores e funcionários da COPIN - UFPB, especialmente a Aninha por sua preciosa dedicação e amizade.

Aos colegas de curso (Aídre, Almir, Gilson, Joseana, Kátia, Robson, Sônia, Vana, Eliane), que tornaram os meus dias de trabalho mais agradável.

Aos companheiros da Embrapa, pelo apoio e dedicação.

Ao professor Francisco de Assis Santos e Silva do DME, pela sua presteza.

A Embrapa e CAPES, por terem subsidiado meus estudos.

As organizações (CANDE, CPQD, CAIXA ECONÔMICA, LMRS, HUMANA, SAELPA, TELPA, TRT, UNIMED), que permitiram a realização da pesquisa. Destacando a CANDE, onde o auxílio do Sr. Alberto Freire dos Santos, foi de fundamental importância.

## *Lista de Figuras*

- Figura 1.1* - Visão simplificada do ambiente de banco de dados, 2
- Figura 1.2* - Arquitetura de ANSI-SPARC, 4
- Figura 1.3* - Componentes de um SGBD, 6
- Figura 2.1* - Nível de satisfação das empresas pesquisadas com relação ao SGBD, 22
- Figura 3.1* - Uma ilustração do ciclo de vida do sistema e do método de decisão LSP, 31
- Figura 3.2* - Árvore do sistema de requisitos e parâmetros de custo (primeiro e segundo níveis): RPC, 28
- Figura 3.3* – Subárvore de implementação de registro, 34
- Figura 3.4* - Uma estrutura genérica de agregação para as três classes de parâmetros, 35
- Figura 4.1* - Fases da metodologia MAPESGBD, 42
- Figura 4.2* - Árvore genérica de requisitos, 44
- Figura 4.3* – Quadro comparativo dos modelos relacional e orientado a objeto., 51
- Figura 4.4* – Representação de relacionamento em um modelo de dados orientado a objetos, 57
- Figura 4.5* – Camadas básicas de ambientes de desenvolvimento de bancos de dados, 60
- Figura 4.6* - Exemplo de refinamento da AGR, 59
- Figura 4.7 (a)* - Representação gráfica da classificação dos grupos de requisitos em classes de requisitos essenciais, opcionais e desejáveis, 61
- Figura 4.7 (b)* - Classificação de grupo de requisitos ambiente de desenvolvimento por classe para o refinamento exemplificado na figura 4.6., 61
- Figura 5.1* - Árvore específica de requisitos, 73



## ***Lista de Tabelas***

*Tabela 2.1* - Perfil das empresas pesquisadas, 12

*Tabela 2.2* - Perfil do ambiente de banco de dados, 15

*Tabela 2.3* - Nível de satisfação das empresas pesquisadas, 17

*Tabela 2.4* – Requisitos desejáveis pelas empresas pesquisadas, 18

*Tabela 2.5* - Conversão de escala, 19

*Tabela 2.6* - Nível de satisfação para o recurso linguagem de consulta na escala semântica, 19

*Tabela 2.7*- Nível de satisfação para o recurso linguagem de consulta na escala numérica, 19

*Tabela 2.8* - Níveis globais de satisfação, 20

*Tabela 4.1* - Lista de índices de satisfação, 66

*Tabela 4.2 (a)* - Lista de produtos baseada no atendimento aos requisitos não quantificáveis, 67

*Tabela 4.2 (b)* - Exemplo de uma tabela de produtos classificados com base na análise de requisitos não quantificáveis, 67

*Tabela 4.3* - Exemplo de uma tabela de produtos classificados com base na análise de desempenho, 67

*Tabela 4.4* – Classificação dos produtos baseada em custos, 69

*Tabela 5.1 (a)* - Lista de requisitos essenciais, 75

*Tabela 5.1 (b)* - Lista de requisitos desejáveis, 75

*Tabela 5.1 (c)* - Lista de requisitos opcionais, 75

*Tabela 5.2* - Lista de produtos classificados por notas, 77

*Tabela 5.3* - Índices de satisfação para os requisitos não quantificáveis, 78

*Tabela 5.4* - Classificação dos produtos baseada na análise dos requisitos Não quantificáveis, 79

*Tabela 5.5* - Planilha de custo dos produtos, 80

## *Lista de Abreviaturas*

**API:** Application Program Interface

**BLOB:** Binary Large Object

**DDL:** Data Definition Language

**DML:** Data Manipulation Language

**JDBC:** *Java Database Connectivity*

**ODBC** *Open Database Connectivity*

**OO:** Orientado a Objeto

**ODMG:** *Object Database Management Group*

**SGBD:** Sistema Gerenciador de Banco de Dados

**SGBDR:** Sistema Gerenciador de Banco de Dados Relacional

**SGBDOO:** Sistema Gerenciador de Banco de Dados Orientado a Objetos

**SO:** Sistema Operacional

**SQL:** *Structure Query Language*

**RAID:** Redundant Array of Independent Disk

**WEB:** *World Wide Web*

## **Sumário**

### **Capítulo 1- Introdução**

- 1.1. Sistemas de Gerência de Banco de Dados, 1
  - 1.1.1. Arquitetura dos Sistemas de Gerência de Banco de Dados, 3
  - 1.1.2. O Ambiente dos Sistemas de Gerência de Banco de Dados , 4
  - 1.1.3 Utilitários para Sistemas de Gerência de Banco de Dados, 7
  - 1.1.4 Facilidades de Comunicação, 7
  - 1.1.5 A Evolução dos Sistemas de Gerência de Banco de Dados, 8
- 1.2. Motivação para a Pesquisa, 9
- 1.3. Organização da Dissertação, 9

### **Capítulo 2 - O Processo de Seleção de SGBDs nas Empresas**

- 2.1 Pesquisa no Mercado Nacional, 11
  - 2.1.1 Universo, Amostra da Pesquisa, 11
  - 2.1.2 Instrumento da Pesquisa, 12
- 2.2 Parte I da Pesquisa – Levantamento do Perfil do Ambiente de Banco de Dados e os Métodos de Escolha Utilizados nas Empresas, 13
  - 2.2.1 Apresentação dos Resultados (Parte I), 13
- 2.3 Parte II da Pesquisa - Níveis de Satisfação com Relação aos Recursos e Características Oferecidos pelo SGBD, 15
  - 2.3.1 Apresentação dos Resultados (Parte II), 17
- 2.4 Análise dos Resultados da Pesquisa, 18
  - 2.4.1 Análise Estatística dos Resultados, 20
- 2.5 Considerações sobre o Processo de Seleção de SGBDs, 22

### **Capítulo 3 - Métodos de Escolha de SGBD**

- 3.1. Abordagens de Seleção de SGBDs Constatadas na Pesquisa, 25
  - 3.1.1. Abordagem Utilizada pela Empresa “B”, 26
  - 3.1.2. Abordagem Utilizada pela Empresa “H”, 27

- 3.1.3. Abordagem Utilizada pela Empresa “T”, 28
- 3.2. Modelo de Decisão para Análise, Comparação e Seleção de SGBD Baseado na Relação Custo Benefício, 30
- 3.3. Seleção a partir de Testes de *Benchmark*, 38
  - 3.3.1 Metodologia de Winconsin, 38
  - 3.3.2. Metodologia de Débito e Crédito, 39
  - 3.3.3. Metodologia de *Benchmark* Especializado, 39
- 3.4. Considerações sobre os Métodos de Seleção, 40

#### **Capítulo 4 – Proposta de Metodologia para Seleção de SGBDs**

- 4.1 Fases da Metodologia, 42
  - 4.1.1 Escolha do *Decision Maker*, 43
  - 4.1.2 Pré-Seleção dos SGBDs, 43
  - 4.1.3 Levantamento e Classificação dos Requisitos, 43
    - 4.1.3.1 Refinamento da Árvore Genérica de Requisitos, 44
      - 4.1.3.1.1 Definição dos Termos da Árvore, 44
    - 4.1.3.2 Elaboração do Questionário, 61
    - 4.1.3.3 Classificação dos Grupos de Requisitos, 61
  - 4.1.4 Classificação dos Produtos, 62
    - 4.1.4.1 Classificação Baseada nas Informações do Fornecedor, 62
    - 4.1.4.2 Classificação Baseada nos Testes de Desempenho, 67
      - 4.1.4.2.1 Análise de Desempenho Baseada na Metodologia de *Benchmark* Especializado, 68
      - 4.1.4.2.2 Análise de Desempenho Baseada em Testes de Tomada de Tempo, 68
    - 4.1.4.3 Classificação Baseada nos Custos, 69
  - 4.1.5 Classificação Baseada na Relação Custo Benefício, 69

#### **Capítulo 5 - Validação da Metodologia Proposta**

- 5.1. Estudo de Caso, 70
  - 5.1.1 Escolha do *Decision Maker*, 70
  - 5.1.2 Pré-Seleção de SGBDs, 71

- 5.1.3 Levantamento e Classificação dos Requisitos, 71
  - 5.1.3.1 Refinamento da Árvore Genérica de Requisitos, 71
  - 5.1.3.2 Elaboração do Questionário, 73
  - 5.1.3.3 Classificação dos Grupos de Requisitos, 74
- 5.1.4 Classificação dos Produtos, 76
  - 5.1.4.1 Classificação Baseada nas Informações do Fornecedor, 76
  - 5.1.4.2 Classificação Baseada nos Testes de Desempenho, 79
  - 5.1.4.3 Classificação Baseada nos Custos, 79
- 5.1.5. Classificação Baseada na Relação Custo Benefício, 80

## *Capítulo 6 – Conclusão*

- 6.1. Considerações Finais, 81
- 6.2. Proposta para Trabalhos Futuros, 82

## *Bibliografia*

*Apêndice I* – Questionário Aplicado às Empresas Usuárias de SGBDs

*Apêndice II* – Questionário Aplicado às Empresas Fornecedoras de SGBDs

## ***Resumo***

A crescente diversidade de Sistemas Gerenciadores de Bancos de Dados - SGBD existentes no mercado, torna a tarefa de escolha de um produto que melhor se adeque às necessidades de uma corporação cada vez mais difícil. Apresentamos neste trabalho a proposta de uma metodologia para auxiliar usuários e empresas no processo de seleção de um SGBD, MAPESGBD- Metodologia de Apoio à Decisão para Seleção de um SGBD. Esta metodologia oferece ao avaliador uma sistemática para o levantamento do perfil das aplicações e do ambiente de banco de dados da organização. A partir do perfil traçado é feita uma avaliação com a consequente classificação dos produtos alternativos em termos: (1) da adequação dos produtos às necessidades da organização, (2) do desempenho do produto, e (3) da avaliação da relação custo-benefício.

***Abstract***

The ever-growing diversity of DBMS products makes hard the task of selecting a product that best fits the needs of an organisation. This work proposes a methodology that aims to be a decision making aid during the process of DBMS selection. The MAPESGBD methodology offers the decision-maker a systematic approach to gather information on the organisation's needs and its profile in terms of its database environment and applications. An evaluation process of the alternative products is then completed, classifying the products in the basis of: (1) product adequacy to the organisation's needs, (2) product performance and, (3) cost-benefit analysis.



## *Capítulo 1*

### ***Introdução***

A informação tem assumido um papel cada vez mais importante no gerenciamento das atividades das organizações, levando a uma crescente utilização de SGBDs como ferramenta de auxílio neste processo.

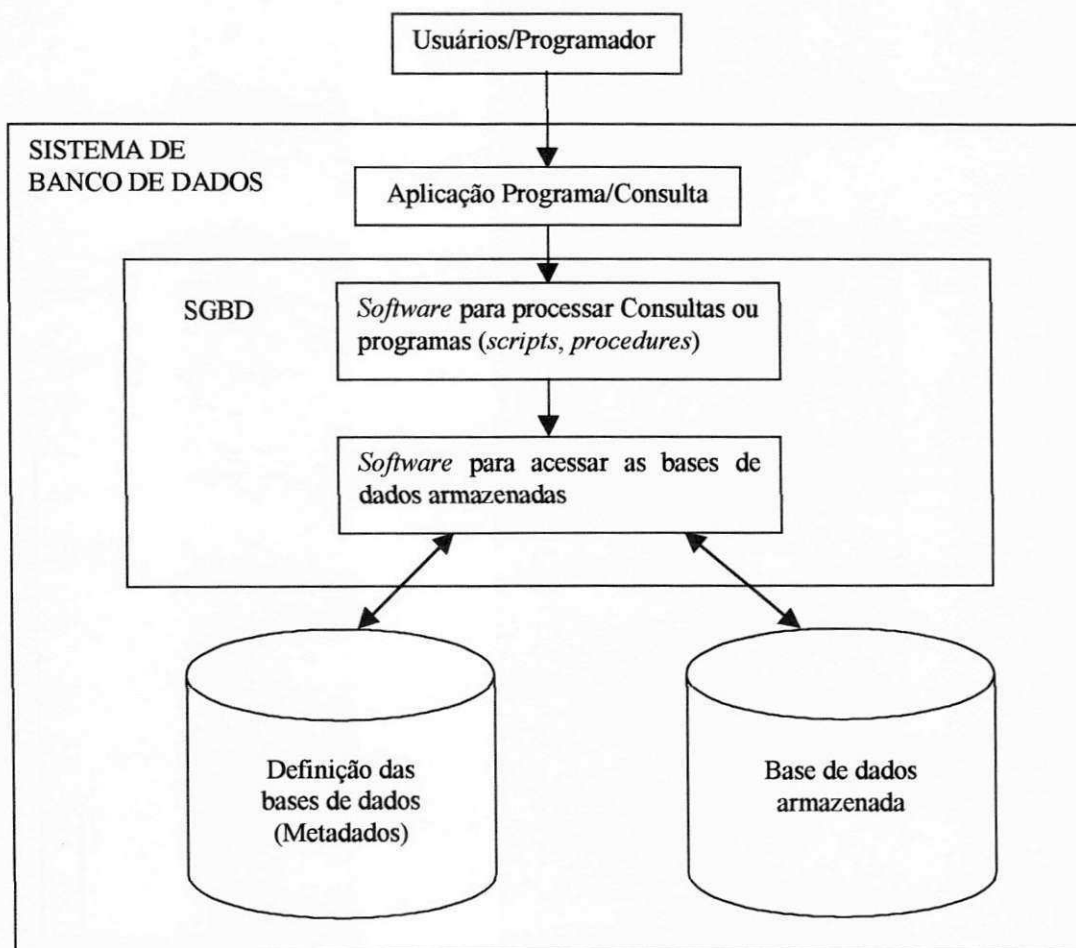
#### **1.1. Sistemas de Gerência de Banco de Dados**

A tecnologia de banco de dados, tem sido um elemento importante para o crescimento do uso de computadores. Portanto, não há exagero ao dizer que os bancos de dados estarão ocupando um espaço importante na maioria das áreas onde os computadores estão sendo usados, incluindo, engenharia, medicina, direito, educação e tantas outras.

Segundo Elmasri & Navathe (1994), um banco dados é uma coleção de dados relacionados (fatos que podem ser registrados e têm um significado implícito). Em um banco de dados encontram-se implícitas algumas características:

- Um banco de dados representa alguns aspectos do mundo real, algumas vezes chamados de “minimundo” ou “universo de discurso”. Mudanças no minimundo refletem-se no banco de dados.
- Um banco de dados é uma coleção lógica de dados coerente, com alguns significados inerentes.
- Um banco de dados é projetado, construído e popularizado com dados de propósito específico, isto é, destinados a um grupo de usuários e a algumas aplicações nas quais os usuários estão interessados.

Um sistema gerenciador de banco de dados (SGBD) é uma coleção de programas que permitem aos seus usuários criar e manter bancos de dados. O SGBD é portanto um aplicativo de propósito geral que facilita o processo de definição, construção e manipulação das bases de dados. A **Definição** das bases de dados envolve a especificação dos tipos de dados, estruturas e as restrições (regras do negócio) para os dados a serem armazenados no banco de dados. A **Construção** das bases de dados é o processo de armazenar dados no SGBD. A **Manipulação** das bases de dados inclui funções para consulta às bases de dados, recuperação de dados específicos e atualização das bases de dados. A Figura 1.1, ilustra uma visão simplificada do ambiente de banco de dados.

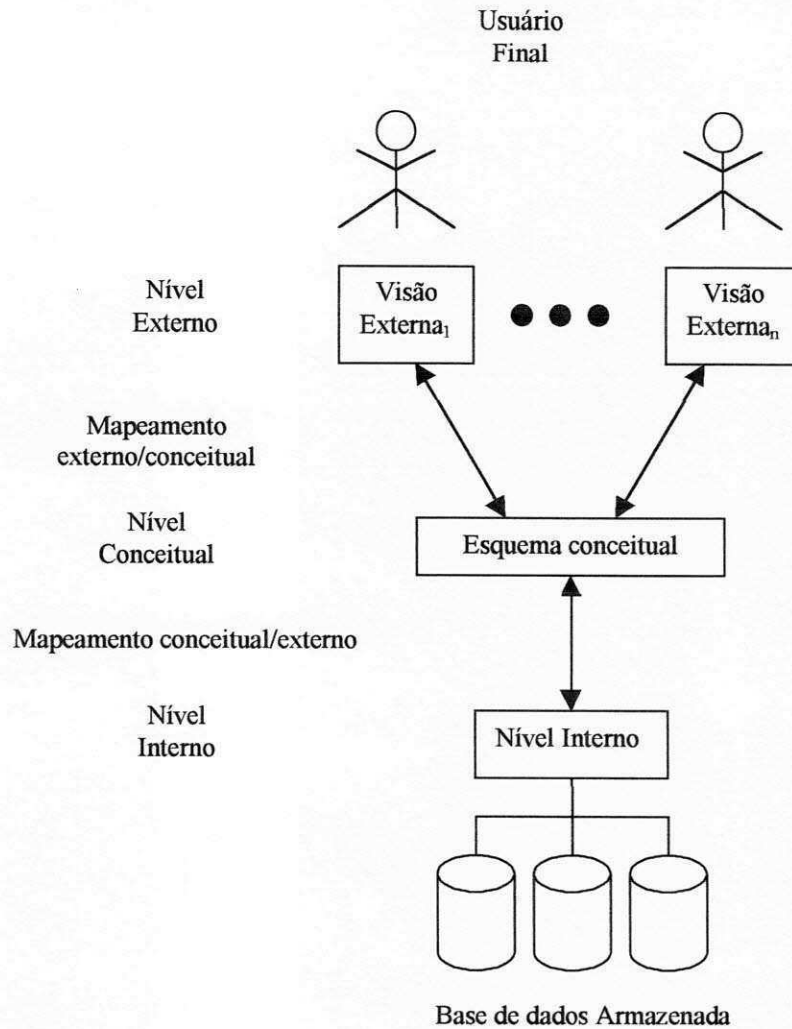


**Figura 1.1** Visão simplificada do ambiente de banco de dados  
Fonte: Elmasri & Navathe (1994)

### 1.1.1. Arquitetura dos Sistemas de Gerência de Banco de Dados

Nesta seção nós discutiremos o modelo ANSI/SPARC, (ANSI, 1986ab), (Elmasri & Navathe,1994) o qual apresenta uma arquitetura específica para banco de dados, denominada arquitetura de três níveis. Uma das principais propostas de um sistema de banco de dados é de prover os usuários com uma visão abstrata dos dados. Isto é, o sistema omite certos detalhes de como os dados são armazenados e mantidos. Entretanto, para que o sistema possa utiliza-los, os dados devem ser recuperados eficientemente. Este conceito tem direcionado o projeto de estruturas de dados complexas para a representação dos dados em bancos de dados. Uma vez que muitos dos usuários de banco de dados não são treinados em computação, a complexidade é superada pela introdução de três níveis de abstração que simplificam a interação do usuário com o sistema.

- O **nível interno** consiste do esquema interno, o qual descreve a estrutura interna (física) de armazenamento dos dados. O esquema interno usa o modelo físico dos dados e descreve em detalhes o caminho de como os dados estão armazenados e como podem ser acessados no banco de dados.
- O **nível conceitual** inclui o esquema conceitual, o qual descreve todas as estruturas do banco de dados para a comunidade de usuários. O esquema conceitual oculta certos detalhes de como os dados estão armazenados fisicamente (esquema interno), apresentando para o usuário apenas os objetos do banco de dados e suas instâncias.
- O **nível externo** inclui o esquema externo ou visão do usuário. Cada esquema externo descreve uma parte do banco de dados que um usuário em particular ou um grupo de usuários estão interessados e oculta o resto do banco de dados daquele grupo de usuários. Um modelo de alto nível ou de implementação pode ser usado neste nível.



**Figura 1.2 Arquitetura ANSI-SPARC**  
**Fonte: Elmasri & Navathe (1994)**

### 1.1.2. O Ambiente dos Sistemas de Gerência de Banco de Dados

O SGBD é um aplicativo complexo. Nesta seção apresentaremos os módulos que o constituem e os tipos de aplicativos com os quais ele interage, ver Figura 1.2. O banco de dados e o catálogo do banco de dados são usualmente armazenados no disco. Os acessos ao disco são controlados, primariamente pelo sistema operacional (SO), o qual escalona a execução das operação de leitura e escrita no disco. O gerenciador de dados armazenados é o módulo responsável por controlar os acessos as informações armazenadas no disco, que fazem parte do banco de dados (instâncias) ou do catálogo (metadados).

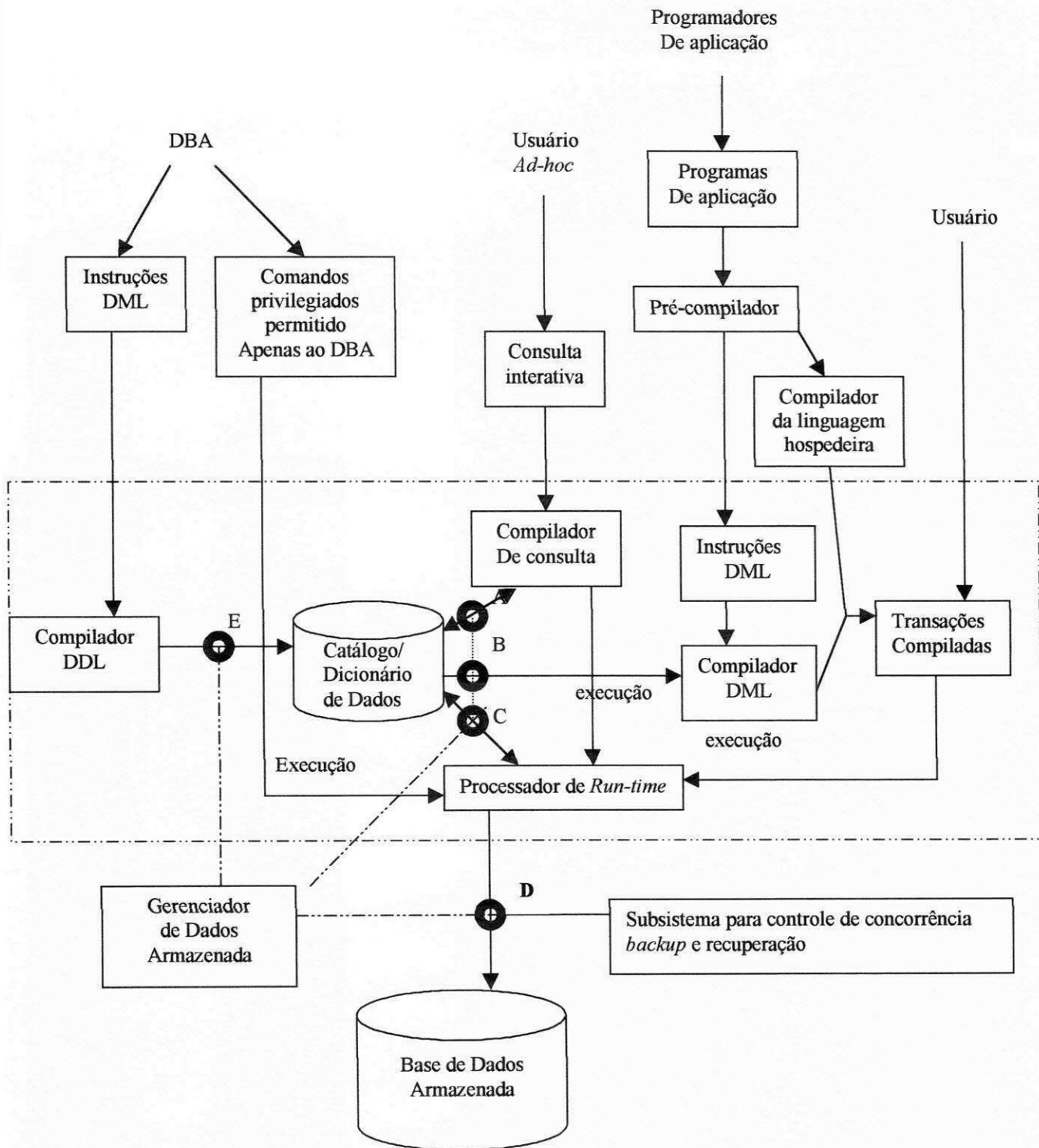
A figura 1.3 detalha o fluxo de tarefas controladas por este módulo. As linhas pontilhadas e círculos marcados pelas letras A, B, C, D e E, representam o fluxo de tarefas

controladas por este módulo. O gerenciador de dados armazenados pode utilizar serviços oferecidos pelo SO para realizar operações de transferências de dados do disco para a memória principal do computador. Uma vez estando no *buffer* da memória principal, os dados podem ser processados pelos outros módulos do SGBD.

O **Compilador DDL** processa as definições do esquema, especificados na DDL, e armazena a descrição do esquema (metadados) no catálogo do SGBD. O catálogo inclui informações como o nome dos arquivos, itens de dados, detalhes de cada arquivo armazenado, mapeamento das informações associadas aos esquemas. O **Compilador de consulta** é responsável por realizar a análise léxica, sintática e gerar chamadas para o processador de *run-time* das consultas que são elaboradas interativamente.

O **pré-compilador** extrai os comandos DML de um programa de aplicação escrito em uma linguagem de programação hospedeira. Estes comandos são enviados para o **compilador DML** para serem compilados e então gerado o código objeto. O restante do programa é enviado para o compilador da linguagem de programação hospedeira. Os códigos objeto gerados pelos compiladores (DML e linguagem de programação hospedeira), são empacotados constituindo uma transação, cujo código executável inclui chamadas para o processador de *run-time*.

O nosso objetivo ao apresentarmos a Figura 1.3, não é o de descrever um SGBD em particular, a nossa intenção é usar esta ilustração para representar módulos típicos de um SGBD. O SGBD interage com o sistema operacional quando acessa o disco para buscar metadados ou instancias do banco de dados. Se o ambiente é multiusuário o SO escalonará as solicitações de acessos a disco e processos de outros usuários. O SGBD também conta com compiladores das linguagens de programação hospedeira. DMLs visuais podem ser oferecidas para facilitar a especificação de consultas ao banco de dados pelos diversos tipos de usuários ilustrados na Figura 1.3.



**Figura 1.3 Componentes de um SGBD**  
 Fonte: Elmasri & Navathe (1994)

### 1.1.3. Utilitários para Sistemas de Gerência de Banco de Dados

Além dos módulos para processamento dos dados descritos anteriormente, muitos SGBDs oferecem utilitários de banco de dados, os quais disponibilizam para o ABD (Administrador de Banco de Dados), um ambiente visual composto por um conjunto de funcionalidades. Apresentaremos a seguir algumas funcionalidades comumente encontradas na maioria desses utilitários.

- **Backup:** Um utilitário de *backup* cria uma cópia do banco de dados. A cópia de *backup* pode ser usada para restaurar em casos de falhas no sistema.
- **Monitoramento de performance:** Este utilitário gera arquivos com estatísticas de performance para o ABD, o qual utiliza estas estatísticas para tomar algumas decisões. Por exemplo, reorganizar os arquivos para melhorar a performance.
- **Dicionário de dados:** Este utilitário gera um arquivo onde são mantidas informações do catálogo sobre os esquemas de banco de dados e restrições, como também outras informações como decisões de projeto, uso de padrões, descrições dos programas de aplicação e outras informações. O dicionário de dados é similar ao catálogo do SGBD, incluindo algumas informações consideradas de grande importância. Segundo Elmasri & Navathe (1994), a combinação de catálogo e dicionário de dados é chamada de **dicionário de dados ativo**. Um dicionário de dados que pode ser acessado apenas pelos usuários e não pelo SGBD é chamado de **dicionário de dados passivo**.

### 1.1.4. Facilidades de Comunicação

O SGBD também precisa interagir com outros programas de comunicação, cuja função é permitir que usuários de sistema de banco de dados, situados em localizações remotas, acessem o banco de dados através de terminais de computador, estações de trabalho, mini ou microcomputadores. Esta comunicação é feita através do uso de linha telefônica, redes de longa distância, ou comunicação por satélite. Muitos SGBDs disponível no mercado incluem em seus módulos pacotes de comunicação. A integração entre os sistemas de banco de dados e os sistemas de comunicação de dados é chamada de sistemas **DB/DC**.

Alguns SGBDs distribuídos são distribuídos fisicamente em muitas máquinas. Neste caso, programas de comunicação são utilizados para conectar as máquinas. O termo **arquitetura cliente/servidor** é usado em conjunção com um SGBD se as aplicações são executadas fisicamente em uma máquina, chamada **cliente** e os dados são armazenados e acessados em outra máquina chamada **servidor**. Várias combinações de cliente e servidor podem ser oferecidas pelo fabricante.

### 1.1.5. A Evolução dos Sistemas de Gerência de Banco de Dados

Os sistemas de bancos de dados baseados nos modelos hierárquico e rede emergiram por volta dos anos 60. Desde de sua introdução em 1970, os gerenciadores de banco de dados baseados no modelo relacional despertaram grande interesse, devido às suas propriedades, incluindo as suas bases formais, homogeneidade, conjunto bem definido de operações algébricas e linguagem baseada no cálculo relacional. Apesar de existirem algumas imperfeições na semântica dos dados do modelo relacional e do expressivo interesse da comunidade de banco de dados e de usuários, por modelos semânticos. A popularidade do modelo relacional tem aumentado de forma considerável. Esta popularidade aumentou com o aparecimento do modelo de entidade e relacionamento (ER) (*Chen, 1976*) modelo hierarquicamente semântico (*Smith & Smith, 1977*) e o modelo de dados semântico (*Hammer & McLeod, 1981*). O modelo relacional estendido, proposto por Codd, (1979) foi sugerido a partir da necessidade de adicionar-se ao modelo relacional certas abstrações proveniente dos modelos anteriores trazendo-lhe um poder mais expressivo em termos de seu modelo semântico. O interesse por modelos semânticos de dados continuou nos anos 80. Hoje, com a expansão dos horizontes das aplicações de banco de dados a necessidade por abstração dos dados tem se tornado cada vez mais relevante.

A forte tendência dos anos 90 para o uso de SGBDs baseados no modelo Orientado a Objeto tem se justificado devido ao fato de que esses sistemas de banco provêm: (1) um poderoso modelo semântico de dados, similar àquele já utilizado nos outros modelos além da incorporação de novos componentes que facilitam a modelagem explícita dos dados, (2) facilidades para implementação de aplicações para banco de dados através da criação de objetos persistente..



## **1.2. Motivação para a Pesquisa**

A seleção de um software consiste de um processo extremamente minucioso, longo e criterioso. Este processo se baseia, entre outros critérios, na avaliação de desempenho, experiências de outros usuários, compatibilidade com o ambiente e a base de software já instalada. São levados em consideração, ainda, a qualidade de suporte técnico oferecida e as condições de atendimento e treinamento.

Ainda como parte essencial do processo de avaliação, os recursos e funcionalidades de cada produto são confrontados com as necessidades concretas do usuário, com a natureza de suas aplicações e com o perfil geral da instalação. Isso porque tais produtos exigem uma plataforma e um ambiente de trabalho.

De uma maneira geral, o conceito de similaridade não se aplica a produtos de software, uma vez que cada produto consiste numa criação intelectual completamente particular, dotada de características próprias, amparada inclusive por direitos autorais. Apesar de diversos produtos serem concebidos com o mesmo propósito, não se pode afirmar que tais produtos sejam igualmente adequados às necessidades do usuário.

Assim sendo este trabalho é sobre a avaliação, comparação e seleção de Sistemas de Gerência de Banco de Dados, propondo uma metodologia de apoio à decisão para a seleção de SGBDs que seja mais adequada à realidade do processo de escolha vivenciado na prática. Como será visto no decorrer deste trabalho, todo o processo de seleção terá como base as necessidades concretas dos usuários, o perfil de suas aplicações e o perfil geral da instalação, visando a aquisição de um produto que ofereça à organização o maior índice de satisfação possível.

Não há pretensões de que esta metodologia conduza a soluções inquestionáveis, uma vez que pela própria natureza heurística do processo de decisão, isto não seria possível. No entanto, ela fornece as sólidas bases para que o SGBD escolhido esteja bem adequado às necessidades de uma organização em particular.

## **1.3. Organização da Dissertação**

Este trabalho está organizado em 6 capítulos, dentre eles este capítulo introdutório. O Capítulo 2, apresenta a pesquisa realizada junto a nove empresas brasileiras, procurando

mostrar como elas se organizam na escolha de seus SGBDs e qual o seu nível de satisfação em relação às escolhas realizadas. O Capítulo 3 tem como objetivo apresentar métodos, critérios e ferramentas de seleção encontrados na literatura e nas empresas pesquisadas. O Capítulo 4 apresenta a proposta de Metodologia para Seleção de SGBDs, objetivo principal deste trabalho. No Capítulo 5, é apresentado um estudo de caso com base na metodologia proposta, de modo a esclarecer sua aplicação. Finalmente, no Capítulo 6 são apresentadas as considerações finais sobre o trabalho e propostas de continuidade.

Os apêndices I e II contêm, respectivamente, o questionário utilizado na pesquisa apresentada no Capítulo 2 e o questionário enviado ao fornecedor durante o estudo de caso apresentado no Capítulo 5.

***O Processo de Seleção de SGBDs nas Empresas.***

O presente capítulo tem como objetivo apresentar a pesquisa realizada recentemente junto a nove empresas brasileiras destacando os métodos e instrumentos utilizados por essas organizações na seleção de SGBDs. O capítulo conclui com a análise dos dados coletados.

**2.1. Pesquisa no Mercado Nacional**

Neste trabalho buscou-se a realização de um estudo exploratório, levando-se em consideração a pesquisa: (1) dos métodos e critérios para seleção de SGBDs nas organizações e (2) do nível de satisfação com relação aos recursos oferecidos pelos mesmos.

A identificação de como o processo de seleção de um SGBD ocorre nas organizações nos permite visualizar quais as dificuldades encontradas para identificar indicadores que possam contribuir no processo de decisão de um SGBD.

Com o propósito de assegurar o sigilo das informações pesquisadas, as empresas são identificadas em um grupo de A a I.

**2.1.1. Universo, Amostra da Pesquisa**

A amostra desta pesquisa consistiu de nove empresas brasileiras, entre estatais e privadas, que atuam em diferentes ramos da sociedade, como pode ser observado na Tabela 2.1. Estas empresas possuem uma característica em comum, utilizam um SGBD para auxiliar na catalogação e na gerência de informações. A consulta foi realizada nas cidades de Campina Grande e João Pessoa no Estado da Paraíba, Campinas no Estado de São Paulo e Brasília no Distrito Federal.

Empresa	Perfil
Empresa A	Uma das seis maiores empresas brasileiras na área de tubos e conexões
Empresa B	Consultora e desenvolvedor de software
Empresa C	Prestadora de serviços de telecomunicações
Empresa D	Prestadora de serviços às cooperativas de serviços médicos
Empresa E	Fornecedora de informações e análises meteorológicas, recursos hídricos e sensoriamento remoto
Empresa F	Prestadora de serviços de eletrificação
Empresa G	Empresa de grande porte com atuação nas áreas financeira e social
Empresa H	Instituição pública ligada à Justiça do Trabalho
Empresa I	Instituição pública ligada à pesquisa

**Tabela 2.1 - Perfil das empresas pesquisadas**

### 2.1.2. Instrumento da Pesquisa

Os dados desta pesquisa foram obtidos através de entrevistas com base em um questionário estruturado apresentado no Apêndice I, conforme sugere em (Boyd & Westfall 1987). Este questionário foi elaborado tendo em vista os objetivos da pesquisa, detalhados a seguir, e sua aplicação foi através de entrevista com os usuários dos SGBDs nas empresas pesquisadas.

A primeira parte do questionário teve como objetivo identificar os métodos de seleção utilizados por estas empresas na escolha de um SGBD; já a segunda parte objetivou identificar o nível de satisfação dos usuários com relação aos recursos e características oferecidos pelo SGBD instalado.

Do questionário constam 28 variáveis de pesquisa, V1 a V28, agrupadas em: Ambiente da Empresa, Aspectos gerais, Ambiente de desenvolvimento, Administração dos dados e Administração do SGBD.

A especificação dos Critérios de Seleção se fundamentou nas recomendações feitas na literatura (Su et al. 1987), (Gradarin & Valdúriez 1989), (Pereira 1990) e no levantamento das características relevantes para cada tipo de aplicação. Esta pesquisa serviu de base, para identificar os critérios utilizados e as dificuldades encontradas na realização do processo de seleção.

## **2.2. Parte I da Pesquisa – Levantamento do Perfil do Ambiente de Banco de Dados e os Métodos de Escolha Utilizados nas Empresas**

Esta parte do questionário teve como objetivo principal identificar os métodos de seleção utilizados pelas organizações para a escolha de um SGBD, através da aplicação de questões subjetivas.

A partir da variável “V1- SGBDs Instalados”, foram colhidas informações sobre o produto mais utilizado na empresa, tais como: ambiente, plataforma e versão. A segunda variável, “V2 - Perfil dos Usuários”, buscou identificar os tipos de usuários de SGBD entrevistados durante a pesquisa, de modo a melhor interpretar suas expectativas e o nível de satisfação quanto ao produto instalado. A classificação utilizada foi aquela proposta por Burns et al. (1986) e Korth & Silberschitz (1993). Segundo estes autores de acordo com o tipo de serviço requerido pelo usuário ao banco de dados, é possível identificar seis categorias de usuários: final *ad-hoc*, paramétrico, programador, projetista, administrador de banco de dados e analista de software básico.

A variável “V3 – Método de seleção”, diz respeito ao método de seleção adotado pelas empresas pesquisadas, na figura do *decision-maker*, responsável pela escolha do produto. Neste trabalho os métodos de seleção foram classificados em: sistemáticos e implícitos. O método sistemático define uma seqüência de etapas a serem seguidas, até se chegar a um produto que possa atender às necessidades da organização, enquanto que no método implícito não existe uma seqüência de etapas bem definidas para se processar à escolha.

A variável V4 - Fontes pesquisadas, representa as fontes de informação consultadas sobre o produto durante o processo de seleção como: material bibliográfico (artigos, periódicos), material de divulgação (folders, manuais), profissionais experientes (pesquisadores da área de banco de dados, administradores de base de dados, projetistas).

### **2.2.1. Apresentação dos Resultados (Parte I)**

Esta seção tem como objetivo apresentar um quadro resumo: (1) do perfil de banco de dados das empresas pesquisadas (subseção 2.1), (2) dos níveis de satisfação dos usuários com relação a recursos e características dos SGBDs das empresas pesquisadas,

quanto aos critérios propostos neste trabalho e as recomendações a serem incorporadas aos recursos ou características, para que estes possam vir a atender as suas necessidades.

Com base no levantamento realizado, ver Tabela 2.2 foi possível identificar que no perfil do ambiente de banco de dados das empresas pesquisadas, existe uma forte predominância por utilização de sistemas de gerência de banco de dados relacional, multiusuário nas plataformas Windows e UNIX.

A distribuição de profissionais de acordo com o tipo de serviço requerido ao banco de dados caracteriza o perfil de usuários do ambiente. Na grande maioria das empresas pesquisadas identificou-se a existência de um **usuário final** o qual tem pouca experiência com a tecnologia de processamento de dados. Este tipo de usuário pode ser classificado em: usuário *ad-hoc*, o qual interage com o sistema através da formulação de consultas através das chamadas LMDs - Linguagens de Manipulação de Dados e usuário *paramétrico*, o qual interage com o sistema evocando programas aplicativos

Em muitos casos, as atividades de **programador de aplicação e projetista** são exercidas pelo mesmo usuário. Estas atividades consistem respectivamente em: (1) interagir com o sistema através de uma linguagem hospedeira na qual são escritos os programas de aplicação e (2) projetar o banco de dados, especificando os esquemas conceitual e externo para que possam ser definidos pelo DBA. Além destes usuários destacamos o **administrador de Banco de Dados** o qual interage com o sistema através das LMD, LDD - Linguagem de Definição de Dados, ferramentas de auditoria, ferramentas de controle de acesso e recursos para recuperação e definição de restrições de integridade. O **analista de software básico**, juntamente com o administrador de banco de dados, interage com o sistema para resolver problemas de desempenho e alocação de espaço em disco e de memória.

Cerca de 50% da amostra selecionou seu sistema de gerência de banco de dados, com base apenas em consultas a material bibliográfico, material de divulgação e profissionais experientes, sem referenciar qualquer metodologia. Aproximadamente 30% (empresas B, H e I) escolheu seu SGBD baseado em um método sistemático de seleção, ver Capítulo 3. Os 20% restante considera-se desconhecido o método utilizado para seleção do produto pois os profissionais responsáveis pelo processo de seleção já não mais se encontram na empresa e não foi encontrada nenhuma documentação a respeito.

Contudo através da análise estatística dos resultados, constatou-se que estes 30% da amostra alcançaram um melhor índice de satisfação com relação ao SGBD selecionado do que os demais.

Pesquisa – Parte 1	Empresas Pesquisadas								
	A	B	C	D	E	F	G	H	I
<b>V1. SGBDs Instalados</b>									
Possui um SGBD Instalado	✓	✓		✓		✓		✓	
Possui dois SGBDs Instalados							✓		
Possui três SGBDs Instalados			✓		✓				✓
SGBD: Relacional	✓	✓	✓	✓	✓	✓	✓	✓	✓
Outros									
Monousuário		✓			✓				
Multiusuário	✓		✓	✓		✓	✓	✓	✓
Plataforma DOS	✓				✓				
Plataforma UNIX	✓	✓		✓				✓	✓
Plataforma VM/ESA						✓			
Plataforma MVS							✓		
<b>V2. Perfil dos Usuários</b>									
Usuário Final ad-hoc		✓					✓		
Usuário Final	✓	✓	✓	✓	✓	✓	✓	✓	✓
Programador									
Projetista									
Administrador de Banco de Dados			✓		✓	✓	✓	✓	✓
Analista de Software Básico			✓		✓	✓	✓	✓	✓
Programador e Projetista	✓	✓	✓	✓	✓	✓	✓	✓	✓
DBA e Analista	✓	✓		✓					
<b>V3. Método de Seleção</b>									
Desconhecido	✓						✓		
Ausente			✓	✓	✓	✓			
Sistemático		✓						✓	✓
<b>V4. Fontes Pesquisadas</b>									
Desconhecido	✓				✓				
Material Bibliográfico		✓					✓	✓	✓
Material de Divulgação		✓	✓	✓		✓		✓	✓
Profissionais Experientes		✓						✓	✓

Tabela 2.2 - Perfil do ambiente de banco de dados

### 2.3. Parte II da Pesquisa – Níveis de Satisfação com Relação aos Recursos e Características Oferecidos pelo SGBD

Esta parte da pesquisa, objetivou identificar o nível de satisfação dos usuários com relação aos recursos e características oferecidos pelo SGBD instalado, através de questões subjetivas e de múltipla escolha.

### ***Aspectos Gerais***

Este grupo de variáveis, definido em (Martins, 1993), procura identificar o nível de satisfação das empresas a partir da opinião dos usuários entrevistados sobre as características do produto. Fazem parte do conjunto de características consideradas relevantes pelos autores: V5. Documentação do produto; V6. Treinamento no uso do produto e V7. Suporte técnico.

### ***Ambiente de Desenvolvimento***

Este grupo de variáveis refere-se aos recursos oferecidos pelo Ambiente de desenvolvimento e são: V8. Linguagem de consulta; V9. Gerador de relatórios; V10. Gerador de formulários; V11. Ferramentas CASE; V12. Linguagem de programação; V13. Adequação ao desenvolvimento das aplicações da empresa e V14. Portabilidade de código.

### ***Administração dos Dados***

Este grupo de variáveis da pesquisa buscou identificar os recursos oferecidos pelos produtos para realizar a administração de dados. A conceituação dessas variáveis pode ser encontrada em (Burns, 1986) (Korth & Silberschtz 1993). As variáveis selecionadas neste contexto são: V15. Dicionário de dados; V16. Recursos para alteração de esquema; V17. Métodos de acesso aos dados; V18. Recursos para otimização de consulta; V19. Recursos para controle de concorrência; V20. Recursos para assegurar a integridade de dados e V21. Recursos para a organização física dos dados.

### ***Administração do SGBD***

Na análise dos recursos oferecidos ao Administrador do Banco de Dados, foi selecionado o seguinte grupo de variáveis de pesquisa, definidos em (Korth & Silberschtz 1993) e [Burns, 86]: V22. Recursos para definição de níveis de segurança de acesso; V23. Mecanismos para recuperação da base de dados; V24. Recursos para auditoria do sistema. Aqui foram também pesquisados aspectos do produto relacionados à tarefa do administrador: V25. Adequação do produto a arquitetura do ambiente instalado (Escalabilidade); V26. Facilidade de Integração do SGBD com o sistema operacional; e V27 Facilidade e apoio no Procedimentos de Instalação.

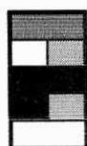


### 2.3.1 Apresentação dos Resultados (Parte II)

Esta seção tem como objetivo: 1) apresentar um quadro resumo dos níveis de satisfação dos usuários de banco de dados das empresas participantes da amostra, mo que diz respeito a alguns recursos e características de SGBD representados pelas variáveis de V5 a V27, ver Tabela 2.4, 2) comentar algumas considerações feitas pelos entrevistados referente a essas variáveis, visando justificar as suas insatisfações, ver Tabela 2.5.

Recurso ou Característica	Nível de Satisfação das Empresas Pesquisadas								
	A	B	C	D	E	F	G	H	I
<b>Aspectos Gerais</b>									
V5. Documentação									
V6. Treinamento									
V7. Suporte Técnico									
<b>Ambiente de Desenvolvimento</b>									
V8. Linguagem de Consulta									
V9. Gerador de Relatório									
V10. Gerador de Formulário									
V11. Ferramentas CASE									
V12. Linguagem de Programação									
V13. Adequação ao desenvolvimento de aplicações Críticas									
V14. Portabilidade de Código									
<b>Administração dos Dados</b>									
V15. Dicionário de Dados									
V16. Recurso para Alterar Esquema									
V17. Métodos de Acesso aos Dados									
V18. Otimizador de Consulta									
V19. Controle de Concorrência									
V20. Integridade de Dados									
V21. Organização Física dos Dados									
<b>Administração do SGBD</b>									
V22. Recursos para Definição dos Níveis de Segurança									
V23. Mecanismos para Recuperação da Base de Dados									
V24. Recursos para Auditoria do Sistema									
V25. Adequação do SGBD a Arquitetura do ambiente									
V26. Interligação do SGBD com o Sistema Operacional									
V27. Procedimentos de Instalação									

Tabela 2.3 - Nível de satisfação das empresas pesquisadas



- Recurso satisfaz as necessidades do usuário
- Recurso oferecido satisfaz em parte as necessidades do usuário
- Recurso oferecido não satisfaz as necessidades do usuário
- Recurso não é oferecido pelo SGBD
- Informação não disponível

Recurso ou Característica	Recomendações
<b>Aspectos Gerais</b>	
V5. Documentação	<ul style="list-style-type: none"> <li>• Uso de língua nacional na confecção dos manuais</li> <li>• Maior clareza nas explicações apresentadas</li> <li>• Maior utilização de exemplos ilustrativos</li> </ul>
V6. Treinamento	<ul style="list-style-type: none"> <li>• Custos adequado à realidade de mercado nacional</li> <li>• Maior treinamento dos monitores.</li> </ul>
V7. Suporte Técnico	<ul style="list-style-type: none"> <li>• Maior domínio da tecnologia</li> <li>• Maior disponibilidade do pessoal técnicos</li> <li>• Redução no tempo de espera pelo atendimento</li> </ul>
<b>Ambiente de Desenvolvimento</b>	
V8. Linguagem de Consulta	<ul style="list-style-type: none"> <li>• Utilizar padrão SQL.</li> <li>• Oferecer ambiente visual</li> </ul>
V9. Gerador de Relatório	<ul style="list-style-type: none"> <li>• Oferecer ambiente visual</li> </ul>
V10. Gerador de Formulário	<ul style="list-style-type: none"> <li>• Oferecer ambiente visual</li> </ul>
V11. Ferramentas CASE	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
V12. Linguagem de Programação	<ul style="list-style-type: none"> <li>• Maior número de recursos de programação</li> <li>• Suportar SQL embutido</li> </ul>
V14. Portabilidade de Código	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
<b>Administração dos Dados</b>	
V15. Dicionário de Dados	<ul style="list-style-type: none"> <li>• Garantir interação consistente e automática entre as informações contidas no dicionário de dados e o esquema de banco de dados.</li> </ul>
V16. Recurso para Alterar Esquema	<ul style="list-style-type: none"> <li>• Oferecer recursos para alteração do esquema sem precisar recriá-lo</li> </ul>
V17. Métodos de Acesso aos Dados	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
V18. Otimizador de Consulta	<ul style="list-style-type: none"> <li>• Oferecer recursos para visualização das estruturas com os respectivos tempos de resposta a uma consulta</li> </ul>
V19. Controle de Concorrência	<ul style="list-style-type: none"> <li>• Disponibilizar recursos para monitoramento de deadlock</li> <li>• Suportar <i>lock</i> de linha</li> </ul>
V20. Integridade de Dados	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
V21. Organização Física dos Dados	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
<b>Administração do SGBD</b>	
V22. Recursos para Definição dos Níveis de Segurança	<ul style="list-style-type: none"> <li>• Inclusão do recurso</li> <li>• Menor complexidade na utilização do recurso.</li> </ul>
V23. Mecanismos para Recuperação da Base de Dados	<ul style="list-style-type: none"> <li>• Oferecer ambiente visual para definição de níveis de segurança</li> </ul>
V24. Recursos para Auditoria do Sistema	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
V25. Adequação do SGBD a Arquitetura do ambiente	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
V26. Interligarão do SGBD com o Sistema Operacional	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>
V27. Procedimentos de Instalação	<ul style="list-style-type: none"> <li>• Nenhuma recomendação citada</li> </ul>

Tabela 2.4 – Requisitos desejáveis pelas empresas pesquisadas

## 2.4. Análise dos Resultados da Pesquisa

Visando uma melhor compreensão dos resultados apresentados, foi feita a conversão da escala semântica dos níveis de satisfação utilizada no questionário, em uma escala numérica de valores, a partir da Tabela 2.5.

Escala Semântica	Escala Numérica
Insatisfatório	0
Regular	1
Bom	2
Excelente	3

Tabela 2.5 - Conversão de escala

A Tabela 2.6 ilustra as respostas (na escala semântica) fornecidas pela empresa A com relação ao recurso Linguagem de Consulta.

Nome do Fator	Excelente	Bom	Regular	Insatisfatório
1. Tempo de Resposta				
2. Interface (Facilidade de Uso)				
3. Qualidade da Apresentação da Informação				
4. Recursos de Hardware e Software				
5. Qualidade de <i>help on-line</i>				

Tabela 2.6 - Nível de satisfação para o recurso linguagem de consulta na escala semântica

A Tabela 2.7 ilustra os dados da Tabela 2.6 convertidos para a escala numérica de valores.

Característica	Nível de Satisfação
Linguagem de Consulta	
Tempo de Resposta	0
Interface (Facilidade de Uso)	0
Qualidade de Apresentação da Informação	0
Recursos de Hardware e Software	0
Qualidade do <i>Help on-line</i>	0

Tabela 2.7- Nível de satisfação para o recurso linguagem de consulta na escala numérica

Em seguida, a partir da expressão 1 foi calculado o nível global de satisfação de cada empresa com relação aos recursos pesquisados.

$$\eta = \sum_{1}^{N} \frac{F_N}{N} \quad (1)$$

Na expressão 1,  $\eta$  representa o nível global de satisfação,  $F_n$  é o fator de pesquisa e  $N$  é o número de recursos pesquisados, constantes da Tabela 2.3. A Tabela 2.8 registra o nível global de satisfação das empresas pesquisadas, com base na expressão 1.

Característica	Nível Global de Satisfação								
	A	B	C	D	E	F	G	H	I
<b>Aspectos Gerais</b>									
V5. Documentação	0	3	2	0	2	2	1.5	1.5	2
V6. Treinamento	0	1.8	0	0	0	1.7	2.7	3	3
V7. Suporte técnico	1	2	2.3	1	0	3	1.3	1.3	1.3
<b>Ambiente de Desenvolvimento</b>									
V8. Linguagem de consulta	0	2.2	1.2	0	1.8	1.8	1.8	2	2.8
V9. Gerador de relatório	2.3	2.7	0	1	3	1.7	1	1.7	1
V10. Gerador de formulário	2.3	2.7	0	1	3	1.7	0	1.7	3
V11. Linguagem de programação	1	3	2	1.5	2	2	2	2.5	2
<b>Administração dos Dados</b>									
V15. Dicionário de dados	0.5	2	0	0	0	0.5	1.5	1.5	2.5
V16. Recursos para alterar esquema	2.3	3	0	2	1	1	2	2	3
V17. Método de acesso aos dados	3	3	0	2	3	2	2	3	3
V18. Otimização de consulta	0	2	0	0	0	2	2	3	3
V19. Controle de concorrência	0	3	0	0	0	2	2	2	3
<b>Administração dos SGBD</b>									
V22. Recurso para definição de níveis de segurança	0	2	0	0	0	2	2.3	3	3
V23. Recurso para recuperação da base de dados	0	3	0	0	0	2	2	2	3
V24. Mecanismo para auditoria do sistema	0	2	0	0	0	2	2	2.5	3

Tabela 2.8 - Níveis globais de satisfação

### 2.4.1 Análise Estatística dos Resultados

Com o objetivo de verificar a existência ou não de uma diferença significativa no nível de satisfação das empresas que utilizaram um método sistemático na seleção de seu produto, com o grupo de empresas que utilizou um método implícito, utilizou-se a Teoria das Pequenas Amostras (Spiegel, 1974), nos níveis de significância 0,01 e 0,05, onde  $n$  número de elementos da amostra e  $n < 30$ . Neste caso  $n=9$ .

Aplicando-se a distribuição de “*Student t*” na análise estatística dos resultados da pesquisa,

se  $\mu_1$  e  $\mu_2$ , representam respectivamente os níveis médios de satisfação do grupo de empresas que utilizaram um método sistemático de seleção para escolher seus SGBDs e do grupo de empresas que utilizaram um método implícito, deve-se decidir entre as hipóteses:

$H_0$ :  $\mu_1 = \mu_2$ , e não há diferença entre os grupos

$H_1$ :  $\mu_1 \neq \mu_2$ , há diferença significativa entre elas

Para a hipótese  $H_0$ :

$$t = \frac{X_1 - X_2}{\sigma \sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \quad (2) \quad \text{onde} \quad \sigma = \sqrt{\frac{N_1 S_1^2 + N_2 S_2^2}{N_1 + N_2 - 2}} \quad (3),$$

$X_1$  = Média aritmética do nível de satisfação das empresas que utilizaram método de seleção;

$X_2$  = Média aritmética do nível de satisfação das empresas que não utilizaram método de seleção;

$N_1$  = Número de amostras do grupo 1;

$N_2$  = Número de amostras do grupo 2;

$S_1$  = Desvio-padrão médio do nível de satisfação das empresas do grupo 1;

$S_2$  = Desvio-padrão médio do nível de satisfação das empresas do grupo 2.

Então:

$$\sigma = \sqrt{\frac{3(0,569)^2 + 6(0,830)^2}{3 + 6 - 2}} = 0,854 \quad \text{e} \quad t = \frac{2,570 - 1,010}{0,854 \sqrt{0,333 + 0,167}} = 2,583.$$

a) Com base em um teste bilateral, no nível de significância 0,01, rejeitar-se-ia  $H_0$ , quando  $t$  estivesse fora do intervalo de  $-t_{0,995}$  a  $t_{0,975}$  o qual, para  $(N_1 + N_2 - 2) = (3 + 6 - 2) = 7$  graus de liberdade, vai de -3,50 a 3,50. Por conseguinte, não se pode rejeitar  $H_0$  no nível de significância 0,01;

b) com base em um teste bilateral no nível de significância 0,05, rejeitar-se-ia  $H_0$ , quando  $t$  estivesse fora do intervalo de  $-t_{0,995}$  a  $t_{0,975}$  o qual, para  $(N_1 + N_2 - 2) = (3 + 6 - 2) = 7$  graus de liberdade, vai de -2,36 a 2,36.

Pode-se, portanto rejeitar  $H_0$  no nível de significância 0,05, concluindo-se que existe diferença significativa entre o nível de satisfação das empresas que utilizaram um método de seleção sistemático com relação às empresas que utilizaram um método implícito.

## 2.5. Considerações sobre o Processo de Seleção de SGBDs

A partir dos dados dispostos na Tabela 8 é possível concluir que as empresas B, H e I, que utilizaram métodos sistemáticos para seleção de seus SGBD, apresentaram níveis globais de satisfação com esses produtos bem superiores àqueles alcançados pelas demais empresas. Estes níveis de satisfação estão representados na Figura 1.

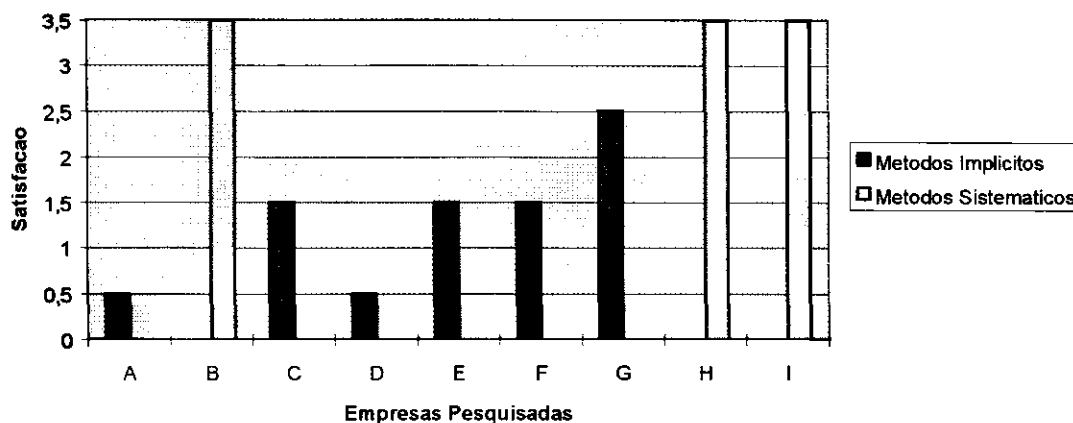


Figura 2.1 - Nível de satisfação das empresas pesquisadas com relação ao SGBD

Analisando-se, as respostas do questionário, também foi possível constatar que a maioria das empresas que adquiriu seus SGBDs com base em métodos implícitos de seleção adquiriu um produto que não possui recursos considerados essenciais, tendo que implementá-los em seus programas de aplicação. Uma outra queixa frequente esteve relacionada às interfaces com o usuário das ferramentas que compõem o SGBD. Estas, na opinião dos usuários, são confusas e pouco amigáveis; levando à subutilização de muitos dos recursos oferecidos. A partir desses resultados, constatou-se a importância da adoção de um método sistemático de seleção.

A adoção de um método sistemático de seleção pode melhorar o nível de satisfação dos usuários com relação ao produto adquirido; isto decorre da investigação das necessidades da organização e da subsequente tradução dessas necessidades em critérios de seleção, para só então ao testar produtos alternativos e verificar quais desses produtos atendem ao maior número de requisitos.

Com base nesta pesquisa também foi possível identificar algumas das razões que levam as empresas a adquirir seus SGBDs sem a utilização de um método sistemático de seleção, os quais passamos a descrever.

- Carência de pessoal especializado disponível para realizar um estudo detalhado das necessidades da organização e, posteriormente, traduzi-las em requisitos desejáveis a um SGBD;
- dificuldade em localizar material técnico capaz de esclarecer os pontos relevantes que devem ser levados em consideração no momento da análise das necessidades da organização;
- carência de métodos de seleção, comparação e avaliação de SGBD capazes de auxiliar a organização de forma prática no processo de escolha;
- pouco conhecimento por parte dos profissionais de informática da empresa, a respeito do vasto universo de conhecimentos relacionados aos recursos oferecidos pelos SGBDs;
- custos associados à alocação de pessoal e recursos necessários ao processo de seleção;
- demandas da própria organização, para que o processo de seleção ocorra em um curto espaço de tempo.

A escolha de um SGBD não é uma tarefa simples e implica uma análise criteriosa das necessidades da empresa. A pesquisa também sugere que a organização disponha de profissionais com conhecimentos na área de banco de dados, para analisar e testar alguns produtos disponíveis no mercado de acordo com as necessidades da organização. A dificuldade reside em como conduzir esse processo de análise se, na maioria dos casos, os profissionais responsáveis pelo processo de escolha pouco vivenciaram tal situação.

As sistemáticas de escolha utilizadas pelas empresas pesquisadas abrangeram desde a seleção baseada em um processo heurístico e não sistemático, onde a principal ferramenta utilizada foi o conhecimento técnico do *decision-maker*, até a definição de uma abordagem sistemática que levou em consideração desde a formação do grupo participante

no processo de seleção até a análise de SGBD, com base em entrevistas junto a fornecedores e usuários.

As abordagens para seleção de SGBD utilizadas pelas empresas pesquisadas basearam-se, principalmente, no levantamento das características que os produtos deveriam apresentar para atender às necessidades da organização. Apenas a empresa H levou em consideração durante a análise, o item custo benefício e nenhuma delas aplicou testes para analisar o desempenho dos produtos, de acordo com o perfil de suas aplicações.

Deve-se acrescentar, a esses procedimentos, que dada à complexidade e à diversidade de sistemas existentes no mercado, torna-se imprescindível a utilização de técnicas de análise de desempenho para a realização de uma avaliação criteriosa do comportamento desses sistemas.

Embora seja possível se encontrar, na literatura metodologias para análise de desempenho de banco de dados que podem ser utilizadas durante o processo de seleção de um SGBD, estas só tornam o processo eficaz se acompanhadas de uma análise criteriosa das características desejáveis do produto, com base nas necessidades da organização.



## ***Métodos de Escolha de SGBD***

Através da pesquisa descrita no Capítulo 2, observou-se que a escolha de um SGBD não é uma tarefa simples e implica uma análise criteriosa das necessidades da empresa. A pesquisa também sugere que a organização disponha de profissionais com conhecimentos na área de banco de dados, para analisar e testar produtos disponíveis no mercado de acordo com as necessidades da organização. A dificuldade reside em como conduzir esse processo de análise, se na maioria dos casos os profissionais responsáveis pelo processo de escolha não vivenciaram tal situação.

A pesquisa realizada evidenciou que dentre as empresas pesquisadas, apenas 33% utilizaram abordagens sistemáticas de escolha durante o processo de aquisição do SGBD, as demais 67% não utilizaram qualquer sistemática de escolha.

Os motivos alegados foram: (1) a indisponibilidade de profissionais com conhecimentos técnicos na área de banco de dados para desenvolver e ou aplicar uma sistemática para seleção, (2) os custos associados a alocação de pessoal e aos recursos necessários para a realização da seleção, e (3) a exigência dos dirigentes das organizações para que o processo ocorra em um curto espaço de tempo.

Neste Capítulo as seções 3.1, 3.2 e 3.3 apresentam respectivamente: (1) as sistemáticas de escolha utilizadas por algumas das empresas pesquisadas, (2) um modelo para comparação, avaliação e seleção de Sistemas de Banco de Dados (Su et al. 1987), e (3) a sistemática baseada na análise de desempenho a partir de testes de *benchmark* (Pereira, 1990).

### **3.1. Abordagens de Seleção de SGBDs Constatadas na Pesquisa**

Nesta seção descreveremos sucintamente as abordagens utilizadas pelas empresas B, H e I para seleção de SGBD, de acordo com a pesquisa.

### 3.1.1. Abordagem Utilizada pela Empresa “B”

A empresa “B” utilizou uma abordagem sistemática de seleção. Nela, a principal ferramenta utilizada foi o conhecimento técnico do *decision-maker*<sup>1</sup>. Por outro lado, as empresas “H” e “I” definiram uma abordagem sistemática levando em consideração desde a formação do grupo participante do processo até a análise de SGBDs com base em entrevistas junto a fornecedores e usuários. No momento de escolher o SGBD o *decision-maker*, baseado na sua experiência profissional e nas necessidades da organização, a empresa percorreu o seguinte caminho:

#### Passo 1 - Levantamento de Características

Neste passo foram levantadas algumas características que deveriam ser encontradas em um SGBD para ser adequado às necessidades de suas aplicações:

1. manipular um grande volume de dados;
2. possuir modelo de dados relacional;
3. assegurar a portabilidade de código para o maior número possível de plataformas;
4. garantir a integridade de dados.

#### Passo 2 - Pré-seleção de SGBDs

Neste momento foram consultados alguns periódicos e revistas especializadas com o objetivo de selecionar alguns SGBDs relacionais disponíveis no mercado para participarem do processo seletivo.

#### Passo 3 - Pesquisa em Empresas

Neste passo foram contatadas algumas empresas que tinham um perfil semelhante ao da empresa “B” para indagar qual o SGBD utilizado e qual o nível de satisfação com relação ao produto.

---

<sup>1</sup> *Decision-maker* = indivíduo, comitê ou gerente da organização, cuja função é selecionar aquele SGBD que melhor se adequa as necessidades da organização.

### 3.1.2. Abordagem Utilizada pela Empresa “H”

Após o levantamento do nível de satisfação dos usuários da empresa com relação ao SGBD instalado, a secretaria de informática da empresa “H” definiu uma sistemática para selecionar o novo produto. Este método consistiu dos seguintes passos.

#### **Passo 1 - Análise do Perfil dos Fornecedores dos Produtos**

Nesta etapa foram selecionados alguns SGBDs relacionais em evidência no mercado (Ingres, Sybase, Informix e Oracle) e convidados alguns de seus fornecedores para:

1. Instalar uma versão do produto para demonstração;
2. Desenvolver um protótipo para ‘o cadastramento de processos’, (aplicação típica da empresa “H”);
3. Proferir palestras e cursos;
4. Promover visitas a usuários do produto.

#### **Passo 2 - Análise dos SGBDs**

Este passo teve como objetivo a definição do conjunto de características requeridas pelo sistema de informação da empresa conforme estão listadas a seguir.

1. Portabilidade;
2. Modelo Relacional e Linguagem SQL;
3. Arquitetura cliente/servidor;
4. Suporte a banco de dados distribuído;
5. Mecanismo de segurança e controle de acesso;
6. Mecanismo de recuperação;
7. Mecanismo de concorrência;

8. Suporte a aplicações multimídia;
9. Dicionário de dados ativo;
10. Ambiente de desenvolvimento com linguagem de quarta geração e gerador de aplicação, telas, menus, e relatórios;
11. Ferramenta para administração de banco de dados;

### **3.1.3. Abordagem Utilizada pela Empresa “I”**

Esta abordagem (TELEBRÁS, 1992) consistiu de cinco de estágios os quais serão descritos a seguir

#### **Estágio 1 - Criação de um Grupo de Trabalho**

Neste estágio foi criado um grupo de trabalho constituído por representantes de oito unidades da corporação, com o objetivo de avaliar SGBDs relacionais, para ambiente UNIX, disponíveis no mercado nacional,.

#### **Estágio 2 - Pré-seleção de SGBDs**

Neste estágio cada unidade participante do processo avaliou e testou integralmente um SGBD. Devido a uma limitação de recursos humanos e computacionais das oito unidades participantes, apenas quatro conseguiram concluir a avaliação, sendo que em duas delas foi possível a avaliação de dois SGBDs. A partir desta etapa foram então pré-selecionados os produtos: *Oracle*, *Informix*, *Ingres*, *Unify* e *Sybase*.

#### **Estágio 3 - Elaboração de uma lista de requisitos a serem analisados**

Com base no levantamento das necessidades da organização, foi elaborada uma lista de requisitos usados como referencial para avaliação dos produtos:.

1. Portabilidade para várias plataformas de hardware/software;
2. Documentação;
3. Suporte técnico;
4. Consumo de recursos;

5. Desempenho;
6. Integridade lógica de dados;
7. DDL: alteração *on-line* de tabelas;
8. Dicionário de dados;
9. Manipulação de cursores em linguagem de terceira e quarta geração;
10. Integridade física dos dados;
11. Controle de concorrência;
12. Controle de acesso;
13. Tolerância a falhas;
14. Métodos de acesso;
15. Arquitetura *multi-threaded*;
16. Ferramentas para administração do BD;
17. Otimização de consultas;
18. Disponibilidade do SGBD;
19. Ambiente de desenvolvimento;
20. Distribuição de processamento;
21. Distribuição de dados.

#### **Estágio 4 - Elaboração de Questionários**

Nesta fase foram elaborados questionários para serem preenchidos pelos fornecedores e usuários dos produtos como parte do processo de avaliação dos SGBDs.

## **Estágio 5 - Apresentação dos Resultados**

Neste estágio foi elaborada uma tabela comparativa entre os SGBDs, apresentando os resultados da avaliação.

### **3.2. Modelo de Decisão para Análise, Comparação e Seleção de SGBD Baseado na Relação Custo Benefício**

O LSP (*Logical Scoring of Performance*) é um modelo de decisão genérico, baseado na relação custo benefício, aplicável para: análise, comparação e seleção de SGBDs (Su et al. 1987). O modelo tem as seguintes características: (1) baseia-se matematicamente em lógica contínua estendida e teoria dos critérios complexos; (2) utiliza procedimentos de decisão genéricos, sistemáticos, bem estruturados e quantitativos; (3) sua técnica se baseia em uma análise de custo e em uma análise de benefícios expressas em termos das preferências do *decision maker*.

A Figura 3.1 ilustra a aplicação do modelo dentro do ciclo de vida do sistema existente. A seguir descrevemos suas etapas.

#### **Passo 1 - Sistema de Requisitos Baseado na Análise de Custos e Parâmetros de Preferência.**

Este passo tem início com a análise das necessidades da organização, no que diz respeito: (1) aos recursos que o SGBD deve oferecer para ser adequado ao perfil do sistema de informação da organização, e (2) aos custos diretos e indiretos associados. Estes custos incluem: o custo do produto em si, da aquisição de seus módulos e de recursos de hardware, software e pessoal técnico para operacionalizar o produto.

Com base nos resultados desta análise é montada uma árvore de Sistema de Requisitos e Parâmetros (SRP). Neste sistema os nós de maior nível representam requisitos, preferências e custos, sendo denominados respectivamente pelas letras 'R', 'P' e 'C', e os nós de menor nível podem representar custo ou preferência apenas. Os nós da árvore SRP podem ser decompostos em um número limitado de filhos para facilitar a associação de pesos, relacionados à importância de cada parâmetro de preferência. Esses

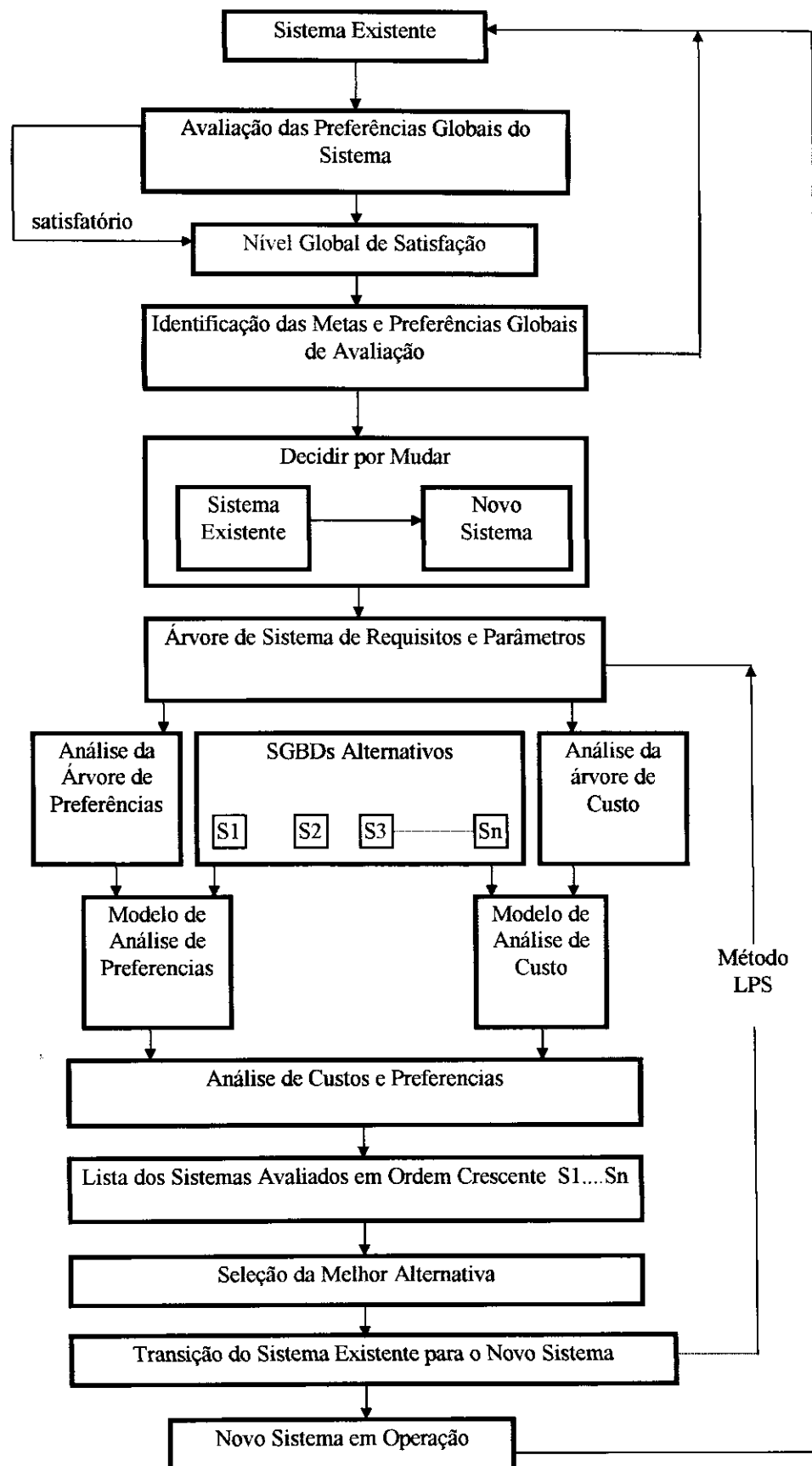


Figura 3.1 - Uma ilustração do ciclo de vida do sistema e do método de decisão LPS

pesos serão utilizados para agregar os nós durante a agregação de parâmetros de preferências (ver Passo 3).

A Figura 3.2 ilustra um exemplo de uma árvore SRP para o exemplo apresentado no artigo de SU em (Su et al. 1987).

### **Passo 2 - Formulação dos Critérios Elementares**

O primeiro passo para análise de requisitos é a formulação dos critérios elementares para os parâmetros de preferência  $(X_1, X_2, \dots, X_n)$  que são os nós folha da árvore de SRP. Um critério elementar é um mapeamento dos valores que podem ser associados a um parâmetro de preferência no intervalo dos reais de zero a um. A associação do valor  $VX_i$  do parâmetro de  $X_i$ , a uma escala de preferência  $E$ , significa que  $E$  expressa o grau de satisfação do *decision-maker* com relação ao parâmetro  $X_i$ , tendo como valor  $VX_i$ .

### **Passo 3 - Agregação de preferências**

A agregação de preferências elementares em preferências globais é o objetivo básico na avaliação dos SGBD sob análise. O processo de agregação se dá através da utilização de funções de agregação de preferência, as quais recebem como entrada as preferências elementares  $e_1, e_2, \dots, e_n$  associadas ao seu peso relativo, e retornam como saída  $E$  as preferências agregadas. Visando facilitar o entendimento do leitor, o autor dividiu este passo da metodologia em três etapas que serão detalhadas a seguir.

**Etapa 1** - Esta etapa consiste na classificação dos parâmetros de preferência das subárvores da árvore SRP dentro de três categorias: essenciais, desejáveis e opcionais, podendo ocorrer situações em que algumas dessas categorias se encontram vazias. Esta classificação tem como base as metas e requisitos da organização com relação ao SGBD a ser adquirido.

Nos parâmetros **essenciais** estão relacionadas as funções que o SGBD precisa exibir para atender aos requisitos daquela subárvore. Os parâmetros **desejáveis** relacionam requisitos extremamente desejáveis, mas não absolutamente essenciais para que os requisitos daquela subárvore sejam satisfeitos. Finalmente, os parâmetros **opcionais** referenciam funções do SGBD que não são essenciais para que os requisitos da subárvore sejam satisfeitos, ou seja, a presença ou ausência de um parâmetro opcional não afeta significativamente a taxa de preferências da subárvore.



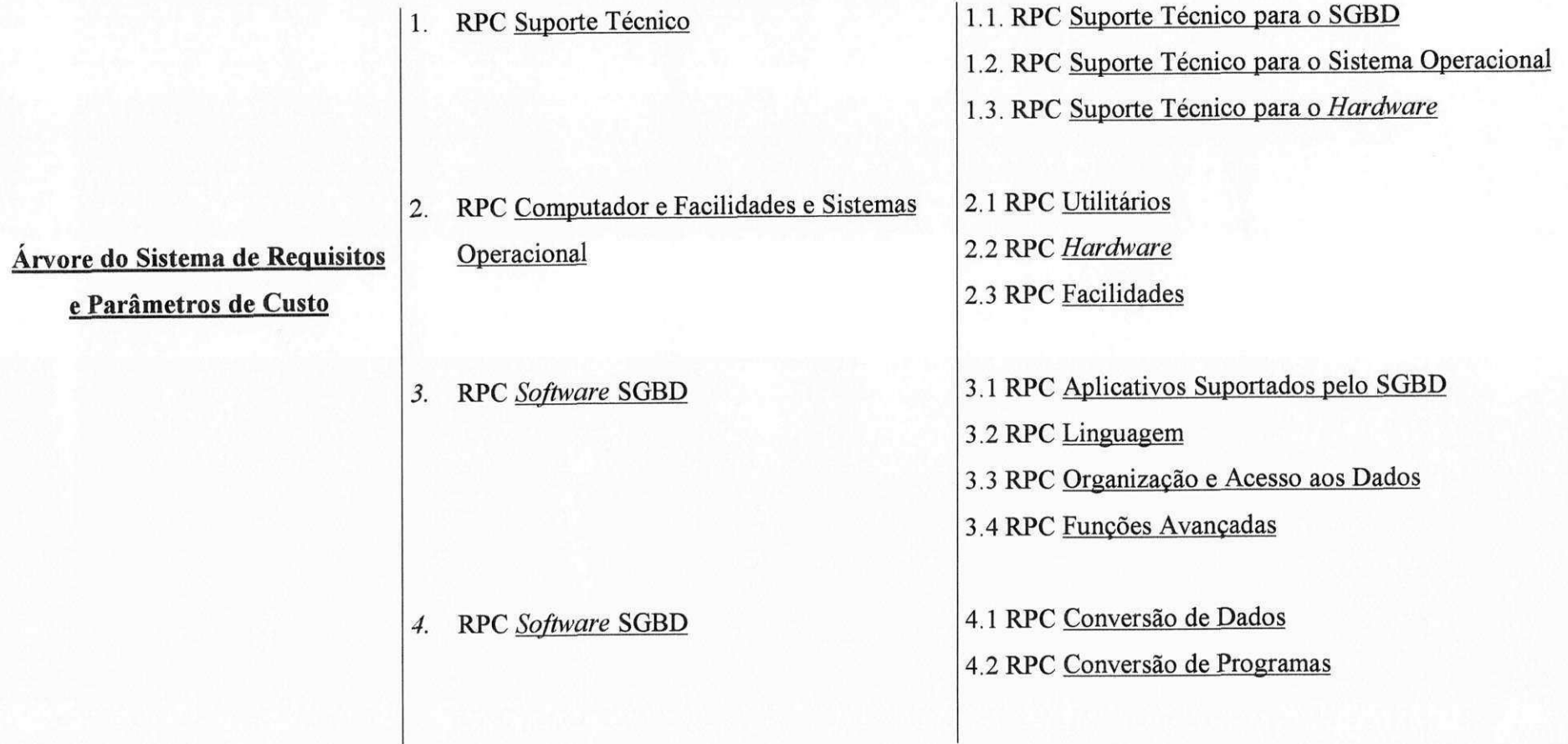
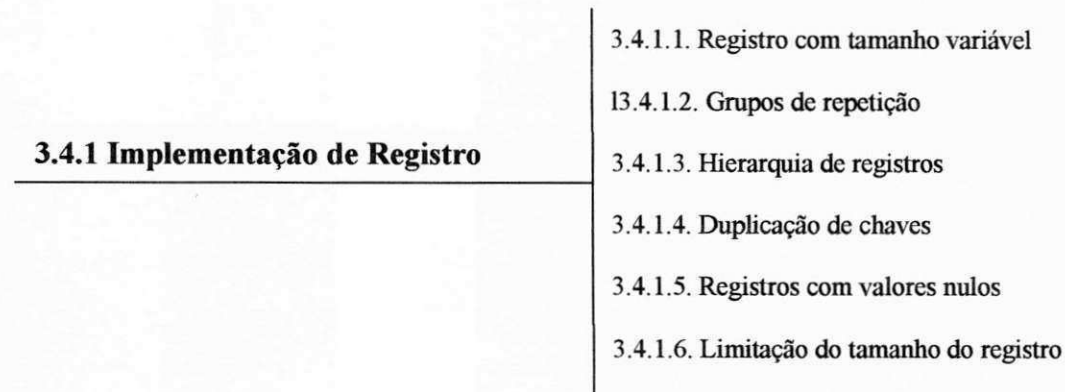


Figura 3.2 - Árvore do sistema de requisitos e parâmetros de custo (primeiro e segundo níveis) RPC  
Fonte: Su et al. (1987)

Como exemplo, considere-se a subárvore "Implementação de registro" pertencente ao nó "organização e acesso aos dados", representada na Figura 3.3. Assim, foi classificado como parâmetro **essencial**: a limitação do tamanho do registro; como **desejável**: registro com tamanho variável e grupos de repetição; e como **opcional**: hierarquia de registros, duplicação de chaves e registros com valores nulos.



**Figura 3.3 - Subárvore de implementação de registro**

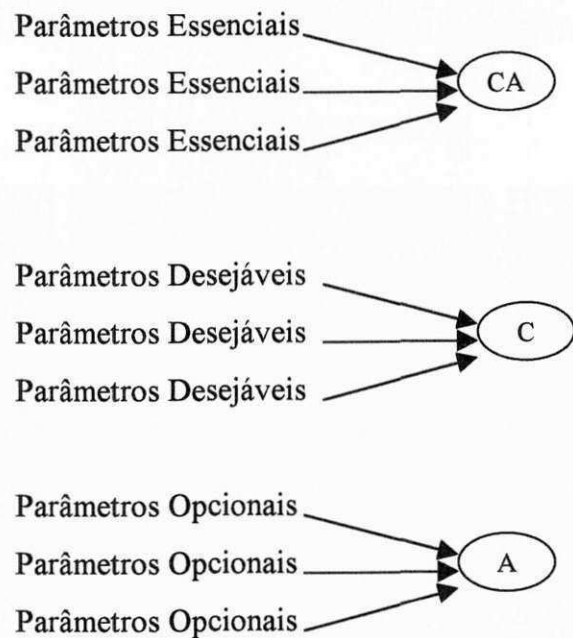
**Etapa 2** - Esta etapa tem como objetivo a agregação dos escores de parâmetros de preferência pertencentes a uma mesma classe, com base na estrutura de agregação ilustrada na Figura 3.4. Como ilustração vejamos o caso seguinte. A obtenção do escore de preferência dos parâmetros opcionais pode ser feita através da média aritmética dos pesos, de acordo com a estrutura de agregação proposta, onde o peso reflete a importância relativa entre os demais parâmetros das classes<sup>2</sup>.

Durante o processo de agregação de preferências pertencentes à mesma categoria de parâmetros, será calculada a média ponderada dos pesos (expressão 3.1), que são as entradas para as funções GCD (função global de disjunção e conjunção)  $CA$ ,  $C$ ,  $A$ ,  $QC$ ,  $QA$ , especificadas na figura 3.4

$$E = (w_1e_1^r + w_2e_2^r + \dots + w_n e_n^r)^{1/r} \quad (3.1)$$

<sup>2</sup> A soma dos pesos para cada classe deverá ser sempre normalizada.

## Agregação de parâmetros dentro de uma classe



## Agregação de parâmetros para as classes básicas

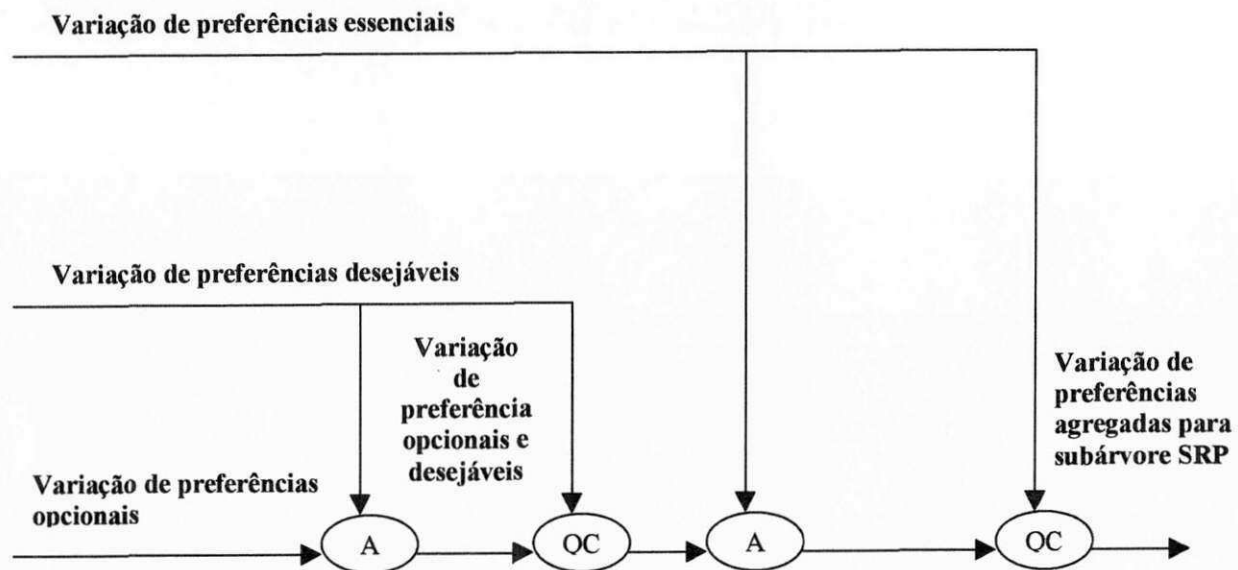


Figura 3.4 - Uma estrutura genérica de agregação para as três classes de parâmetros  
 Fonte: Su et al. (1987)

Em um segundo momento serão calculadas as saídas das funções GCD de cada passo intermediário do processo de agregação, com base nas expressões 3.2 e 3.3, onde  $wr$  será indicado pela GCD. Se a função representada por  $wr$  for uma conjunção, o valor de  $wr$  será o menor valor entre os pesos de entrada para a classe. Se por outro lado for uma disjunção, então  $wr$  será o maior valor e, finalmente se for a função  $A$  (aritmética)  $wr$  será a média aritmética entre os pesos.

$$W_1 = E - wr\delta \quad (3.2)$$

$$W_2 = E - wr\delta^+ \quad (3.3)$$

**Etapa 3** - Esta etapa consiste na agregação das variações de preferências de classes diferentes. Isto será feito incrementalmente pela primeira agregação das variações de preferências das categorias opcionais e desejáveis, e então a agregação do resultado com a variação de preferências essenciais através das funções de absorção parcial. Uma função de absorção parcial pode ser obtida através do cascadeamento de duas funções GCD.

#### **Passo 4 - Análise de Sensibilidade**

A análise de sensibilidade neste contexto é uma investigação sistemática do impacto das funções GCD, dos parâmetros e preferências elementares, na variação das preferências globais. Esta análise é feita com o propósito de identificar aqueles fatores (dentre funções GCD, pesos e preferências elementares), que exercem a maior de influência na variação global de preferências, e portanto quais valores deveriam ser escolhidos para prover uma interpretação significativa dos escores de preferências agregadas.

#### **Passo 5 - Computação e Agregação de Custos**

Este passo é equivalente à análise e agregação de preferências, e pode ser conduzido concorrentemente (de forma independente). A seguir são apresentados os passos para avaliação de custo, sugeridos pelo autor.

- Incluir na avaliação, todos os custos diretos e indiretos;

- Considerar a influência tecnologia sobre complexidade das aplicações, linguagens e escolha de *hardware* e *software* na estimativa de custo de desenvolvimento de *software*;
- Refletir níveis de habilidade pessoal, organizacional/pessoal e o tempo disponível para ser utilizado na estimativa de custo;
- Medir a estimativa para a estrutura da organização, compatibilidade, limitações e outras implicações.

A seguir são apresentadas algumas questões envolvendo estimativa e agregação de custo.

### ***Computação de Custo Usando Módulos de Custo***

Ao construir a árvore de custo, verifica-se que a computação de custo de alguns dos nós é exatamente a mesma, apesar de representarem requisitos diferentes. Visando evitar que o mesmo processo seja executado mais de uma vez, desnecessariamente, e para assegurar que sejam incluídos os mesmos fatores na estimativa dos custos desses nós, sugere-se a criação de módulos custo. Estes módulos podem então ser referenciados sempre que os mesmos parâmetros apareçam na árvore de custo.

### ***O Problema de Agregação de Custo***

A agregação de custos consiste em construir para cada SGBD alternativo um vetor de custos contendo (como em 3.3) os custos para cada requisito, e que leve em consideração a estimativa de vida útil do recurso.

$$C_1, C_2, \dots, C_t \quad \text{onde } t = \text{quantidade de anos} \quad (3.3)$$

Com base no vetor de custos e na taxa de atualização monetária  $r$  será calculado o custo global de cada sistema alternativo, através da expressão 3.4.

$$GC = \frac{C_1}{(1+r)^1} + \frac{C_2}{(1+r)^2} + \dots + \frac{C_t}{(1+r)^t} \quad (3.4)$$

onde  $C$  é o custo anual do produto,  $r$  é a taxa de atualização monetária e  $t$  é o total de anos, que corresponde a vida útil do sistema.

A aplicação do modelo desenvolvido por Su em (Su et al. 1987), apesar de sua formulação bastante abrangente, apresenta algumas limitações, tais como: (1) O modelo é de difícil aplicação, pois utiliza formalismos matemáticos, tais como o cálculo da agregação das preferências, que ficam distantes do avaliador, tanto do ponto de vista de sua formação quanto da sua rotina de atividades. (2) A metodologia não disponibiliza uma ferramenta para automatização dos cálculos, dificultando assim a sua aplicação. (3) Durante a avaliação dos produtos não foi sugerida a aplicação de testes para análise de desempenho dos produtos.

Dada a complexidade e à diversidade de sistemas existentes no mercado, torna-se imprescindível a utilização das técnicas para a realização de uma avaliação criteriosa do comportamento destes sistemas.

A seção seguinte passa a descrever uma metodologia que se baseia na análise do desempenho dos produtos sob teste.

### 3.3. Seleção a partir de Testes de *Benchmark*

Dentre as técnicas disponíveis para análise de desempenho, a técnica de *benchmark* de banco de dados tem sido uma das mais utilizadas. Esta técnica é utilizada, inclusive, na determinação de padrões de avaliação de desempenho de SGBDs, por apresentar resultados com um razoável grau de confiança (Pereira, 1990).

Dentre as metodologias de *benchmark* encontradas na literatura, duas delas têm se destacado devido à sua popularidade, a metodologia Wisconsin utilizada na análise de SGBDs relacionais (Botton et al. 1983) e o *benchmark* de Débito e Crédito utilizado na análise de sistemas de transação on-line (OLTP)<sup>3</sup> (Anderson et al. 1985).

#### 3.3.1. Metodologia de Wisconsin

Também conhecido como *benchmark* de Wisconsin ou DeWitt, este teste tem como objetivo a obtenção de dados de desempenho do SGBD que sejam independentes da escolha de uma aplicação. Caracteriza-se pela utilização de um banco de dados, de um

---

<sup>3</sup> OLTP: *On-line Transaction Processing*

conjunto de transações sistemáticas (ou fictícias) e de um ambiente de execução mono-usuário e *stand-alone*<sup>4</sup>. Suas principais vantagens são: a sua simplicidade, portabilidade e o pouco tempo necessário para a sua execução.

### 3.3.2. Metodologia de Débito e Crédito

Um dos principais objetivos da metodologia de débito e crédito é a obtenção de um padrão de transação para classificar o nível de desempenho de sistemas comerciais tipo OLTP. Esta metodologia é constituída basicamente por três testes: 1) TP1 que consiste na execução de transações concorrentes encontradas em aplicações bancárias, do tipo depósito e saque em conta corrente, 2) teste de busca e atualização sequencial de registros e, 3) teste de ordenação de um grande número de registros.

Recentemente, o TPC<sup>5</sup>, conselho formado pelos principais fabricantes de computadores de grande porte, de estações de trabalho e fornecedores de SGBDs, lançou uma proposta de padronização de *benchmark* baseado no *benchmark* de débito e crédito. A seguir, citaremos alguns dos testes de *benchmark*, encontrados na literatura, que foram padronizados por este conselho.

- **TPC-A** avalia o desempenho dos componentes do sistema em termos do *hardware* e *software* necessários para executar tarefas do tipo OLTP, enfatizando os processos de atualização intensiva do banco de dados (TPC, 1996).
- **TPC-B** destinado a realização da análise de desempenho de aplicações comerciais executadas em *batch*<sup>6</sup>, em ambiente *stand-alone* ou cliente-servidor [Han96].
- **TPC-C** a principal diferença ente o TPC-A e o TPC-C é que o TPC-C trabalha com vários tipos de transações, com bancos de dados mais complexos e com a estrutura global de execução (TPC, 1996).

### 3.3.3. Metodologia de *Benchmark* Especializado

Segundo Pereira em (Pereira, 1990) “A Metrologia de *Benchmark* Especializado MBE tem como principal objetivo tornar-se uma ferramenta de suporte para a análise de desempenho de sistemas de gerência de bancos de dados, podendo modelar convenientemente o ambiente operacional do usuário”.

<sup>4</sup> *Stand-alone*: ambiente em que não existe o compartilhamento de hardware

<sup>5</sup> TPC: *Transaction Performance Council*

<sup>6</sup> *Batch*: aplicações executadas em lotes

### ***Proposta de Metodologia para Seleção de SGBDs***

Muitos critérios devem ser considerados quando se seleciona um SGBD. É preciso lembrar que uma decisão desse tipo é de longo prazo. Não se deve tomar essa decisão sem uma investigação e uma avaliação significativa. É preciso compreender as próprias necessidades de gerenciamento de dados, assim como suas prioridades, antes de tentar comparar produtos concorrentes. O melhor produto é algo que só faz sentido ser considerado nos termos de uma lista, por ordem de prioridade, dos requisitos que o usuário espera que o produto atenda.

Dada a complexidade, a ampla gama de recursos e a diversidade de sistemas existentes no mercado, constatou-se, através da pesquisa apresentada no Capítulo 2, que as empresas que não dispõem de pessoal na área de banco de dados geralmente selecionam os seus SGBDs com base principalmente em parâmetros de custo, tendo como resultado em sua maioria, um baixo índice de satisfação com relação ao produto adquirido.

Com base nos fatores citados anteriormente, a seleção de software não é uma tarefa simples, pois envolve parâmetros heurísticos difíceis de serem quantificados. Inspirado nas sistemáticas de seleção utilizadas por nove empresas brasileiras (ver Capítulo 3), no Modelo de Decisão Baseado em Custo Benefício (LSP) (Su et al. 1987) e na Metodologia de *Bechmark* Especializado (MBE) (Pereira, 1990), descritas no Capítulo 3, desenvolveu-se a MAPESGBD (Metodologia de Apoio à Decisão para Seleção de SGBD).

A Metodologia para Seleção de SGBD, aqui descrita, estabelece uma sequência de cinco etapas descritas na seção 4.1 e tem como principais objetivos: (1) apoiar o processo de seleção de SGBD através da execução de uma sequência de passos preestabelecidos, (2)



reduzir os custos relacionados com a tomada de decisão, e (3) auxiliar as empresas no processo de seleção de um SGBD de forma mais consciente, baseado em parâmetros de custo benefício

#### 4.1. Fases da Metodologia

A Metodologia de Apoio à Decisão para Seleção de SGBD (MAPESGBD) consiste de seis fases (Figura 4.1). Na primeira fase é feita a seleção do indivíduo ou do grupo de indivíduos responsável pelo processo. Em seguida são realizadas as fases de: pré-seleção dos SGBD candidatos a participarem do processo; levantamento e classificação dos requisitos sugeridos pela empresa, de acordo com o perfil de suas aplicações; classificação dos produtos e análise custo benefício. Todas estas fases são interativas e incrementais. A última fase é a apresentação à diretoria da empresa, de um quadro com os resultados de cada fase, concluindo com a indicação do SGBD selecionado.

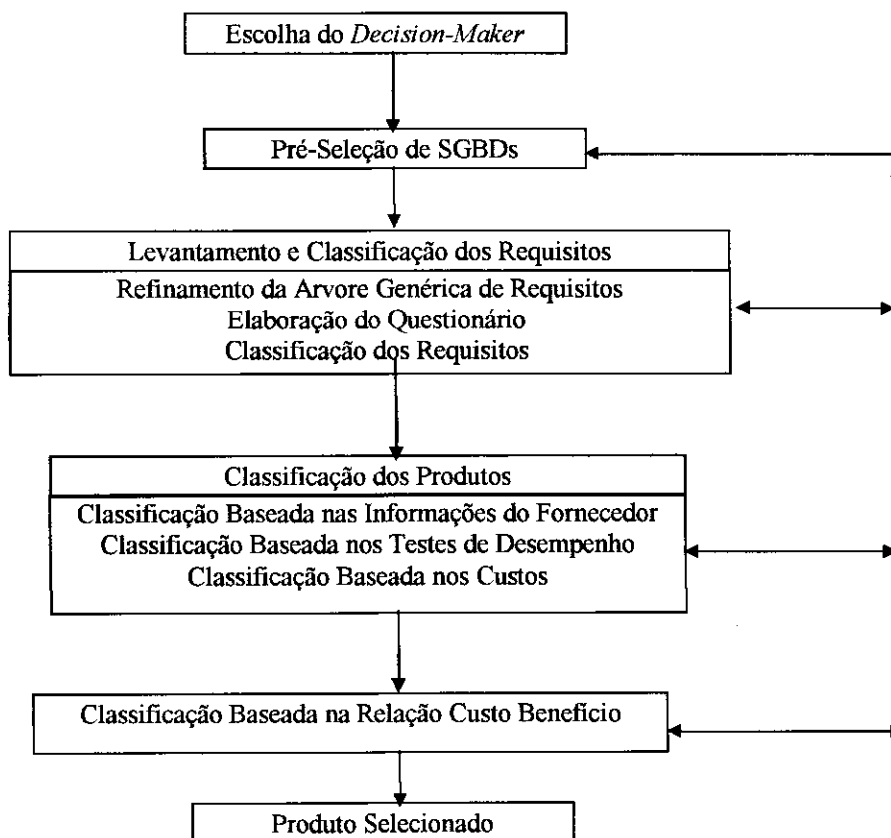


Figura 4.1 - Fases da metodologia MAPESGBD

#### 4.1.1. Escolha do *Decision-Maker*

Esta fase consiste na escolha do indivíduo ou do grupo de indivíduos da organização responsável pelo processo de seleção do SGBD. Sugere-se que o *decision-maker* seja um membro da corporação, que detenha um bom conhecimento da tecnologia de banco de dados e que conheça a organização, para assim poder especificar com maior propriedade quais os requisitos que o produto deverá atender.

#### 4.1.2. Pré-Seleção dos SGBDs

Seleção de alguns produtos disponíveis no mercado com base: (a) nas características do ambiente disponível (recursos de hardware e sistema de operacional) e (b) em consultas a revistas especializadas e usuários de SGBD. Os produtos pré-selecionados compõem o quadro de produtos que participarão do processo seletivo.

#### 4.1.3. Levantamento e Classificação dos Requisitos

Nesta fase metodologia tem como objetivos: (1) realizar um estudo detalhado da empresa, definindo os requisitos em termos de características e recursos que o produto deverá apresentar para atender ao perfil de suas aplicações, (2) refinar a Árvore Genérica de Requisitos (AGR) com base nesse levantamento, produzindo uma Árvore Específica de Requisitos (AER), (3) elaborar um questionário para os fornecedores e (4) classificar os requisitos contidos na AER em: opcionais, desejáveis e essenciais. A seguir descreveremos como alcançar estes objetivos.

A metodologia se baseia em uma árvore Genérica de Requisitos (AGR), representada na Figura 4.2, a qual tem por objetivo apresentar ao *decision-maker* um conjunto de características e recursos básicos necessários a um SGBD. Estas características foram compiladas a partir de: (1) levantamento feito na pesquisa descrita no Capítulo 2, (2) consulta a revistas especializadas, manuais, *folders* de produtos, artigos e periódicos.

O nó raiz da árvore AGR é desmembrado em quatro grupos de requisitos: aspectos gerais, ambiente de desenvolvimento, administração dos dados e administração do SGBD. Não é objetivo desta metodologia cobrir todos os requisitos até os níveis das folhas, uma

vez que estas informações têm um caráter específico de acordo com as características da empresa e de suas aplicações.

#### 4.1.3.1. Refinamento da Árvore Genérica de Requisitos

O refinamento da árvore AGR consiste em acrescentar ou retirar ramificações dos grupos de requisitos da AGR, transformando-a numa AER (Árvore Específica de Requisitos), a partir da especificação dos requisitos necessários ao gerenciador de dados a ser adquirido. Estes requisitos são obtidos a partir do estudo detalhado das aplicações e do ambiente de hardware e software ao qual se destina o SGBD.

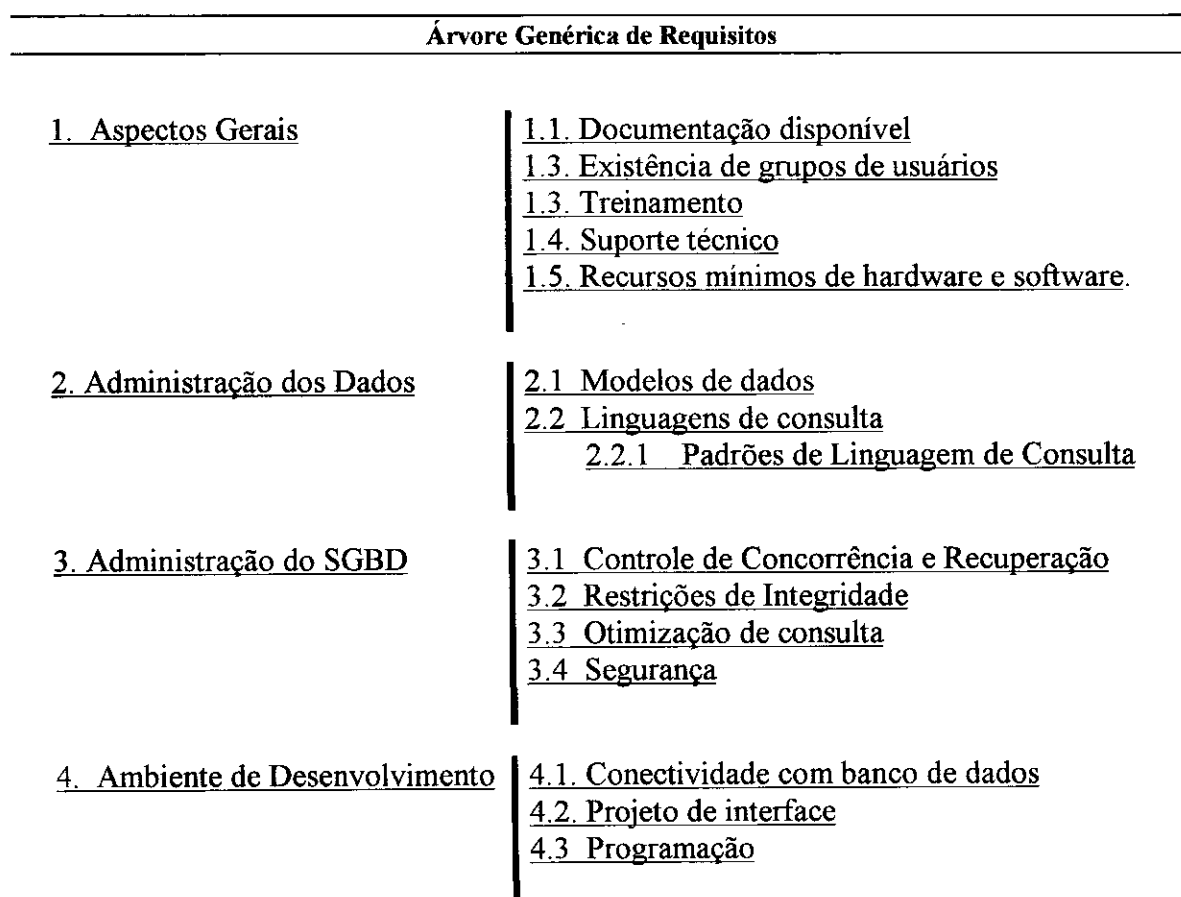


Figura 4.2 - Árvore genérica de requisitos

##### 4.1.3.1.1. Definição dos Termos da Árvore

Nesta seção é apresentada ao leitor uma definição para cada um dos termos presentes na Árvore Genérica de requisitos.

## 1. Aspectos Gerais

A qualidade de um aplicativo, não pode ser avaliada, apenas observando-se o seu conjunto de recursos computacionais. Pontos como, a documentação, disponível, existência de grupos de usuários, treinamento, suporte técnico e as características ergonômicas do ambiente, indispensáveis para o aprendizado, manuseio e obtenção de resultados eficazes em um curto espaço, devem ser observados cuidadosamente, tendo em vista que esta tem sido uma das grandes exigências do mercado.

Os recursos mínimos de hardware e software são também considerados de grande importância, pois aqui serão especificados pelo fabricante a configuração “mínima” e “ideal” para instalação e operacionalização do produto.

### 1.1. Documentação disponível

**Documentação** conjunto de manuais elaborados através de textos explicativos armazenados em papel, CDROM, disquete, fita DAT ou no próprio sistema (tutorial *on-line*) cuja função é auxiliar o usuário na operacionalização do SGBD (Matins, 19r93). Algumas requisitos de uma boa documentação:

- texto explicativo elaborado em português;
- existência de ilustração e exemplos sempre que possível para facilitar o entendimento.

### 1.2 Existência de grupos de usuários

**Grupos de usuários** Este grupo constitui-se de técnicos usuários do produto no Brasil ou no exterior, gerenciado pelo suporte técnico representante do produto, com o objetivo de realizar discussões sobre o produto através da mídia (internet, telefone ou carta), nos seguintes casos:

- dúvidas na operacionalização do produto;
- divulgação de novas versões e inovações lançadas;

- divulgação de eventos (congressos, simpósios, treinamentos) sobre o produto;

submissão de sugestões para correções de possíveis erros encontrados na implementação do produto ou incorporação de novos módulos ao produto, com o objetivo de atender às necessidades da comunidade de usuários

### 1.3. Treinamento

**Treinamento** refere-se às opções de cursos oferecidos pelo fabricante com o objetivo de familiarizar os usuários com os novos conceitos trazidos pelo produto como também no entendimento da melhor forma de operacionalizar o SGBD, utilizando eficientemente os recursos (Martins, 1993).

### 1.4. Suporte técnico

**Suporte técnico** refere-se às atividades de suporte a um produto de software. Justifica-se na busca contínua de satisfação plena do usuário, garantindo que suas dificuldades sejam resolvidas no menor espaço de tempo possível, possibilitando a utilização ótima dos recursos oferecidos pelo SGBD (Martins, 1993).

### 1.5 Recursos mínimos de hardware e software

**Recursos mínimos de hardware e software** refere-se às configurações mínima e ideal de recursos como: memória principal, CPU, disco rígido, sistema operacional, sugerida pelo fabricante para instalação e operacionalização do SGBD, nos ambientes de desenvolvimento e produção.

Segundo (Elmasri & Navathe, 1994), nos últimos trinta anos tem havido uma grande revolução no crescimento da capacidade de armazenamento dos computadores, miniaturização dos circuitos e aumento na velocidade processamento. O custo dos componentes tem caído constantemente. O volume de dados que pode ser processado em um certo espaço de tempo tem crescido. Mais do que nunca grandes bases de dados estão sendo instaladas em pequenos computadores. Até recentemente as grandes bases de dados só podiam ser manuseados por *mainframes* ou minicomputadores. Agora com o aparecimento de máquinas mais rápidas é possível instalar grandes SGBDs em estações de trabalhos e computadores pessoais

Embora o custo de armazenamento tenha diminuído bastante e a capacidade de armazenamento da memória principal dos sistemas tenha crescido consideravelmente, O processamento global dos sistemas de banco de dados é ainda uma grande preocupação. Para resolver estes problemas, muitas sugestões alternativas tem sido feitas para construção de *hardwarees* especiais preparados para as funções de gerenciamento de dados. Estas alternativas, genericamente conhecidas como **máquina de banco de dados** incluem processadores de *back-end*, dispositivos inteligentes, multiprocessadores, memória associativa e o processador de propósito geral.

## 2. Administração dos dados

Se não é necessária a existência de recursos mais elaborados tais como: controle de concorrência, ou seja, se não está prevista a ocorrência de acessos concorrentes aos dados através de operações de leitura ou escrita, se não são feitas consultas *ad-hoc*, ou as aplicações não trabalham com grandes quantidades de dados, então provavelmente não é necessário a aquisição de um sistema de gerência de banco de dados para atender as necessidades de suas aplicações. O desenvolvimento de uma aplicação em uma linguagem de programação convencional, acessando dados estruturados em arquivos pode ser a melhor solução.

Se as operações que irão manipular os dados podem ser expressas em SQL, e as bases de dados podem ser representados de forma tabular com modestas extensões como BLOBs e *procedures* armazenadas ou se estão previstos acessos aos dados através de transações *on-line* então um dos SGBDs com modelo de dados relacional disponíveis no mercado pode ser a melhor escolha.

Se outros requisitos mais especiais são requeridos ao modelo de dados então uma solução possível seria um gerenciador baseado no modelo relacional estendido.

Se for requerida uma completa integração com a linguagem de programação ou se suas aplicações verificam objetos que são usados por mais de que poucos segundos, sendo a aplicação da filosofia de objetos persistentes é insatisfatório então a utilização de um SGBD com modelo de dados orientado a objeto deveria ser considerado.

Segundo (Elmasri & Navathe, 1994), uma característica fundamental da abordagem de banco de dados é que ele fornece alguns níveis de abstração de dados pela omissão de

detalhes de como os dados são por ele armazenados, o que é interessante para muitos usuários de banco de dados. Um modelo de dados é a principal ferramenta para o provimento desta abstração.

## 2.1 Modelo de dados

Um modelo de dados é uma coleção de ferramentas conceituais que são utilizadas para descrever a estrutura de um banco de dados, os relacionamentos entre os dados, a semântica dos dados e restrições de consistência. Dentre os vários modelos de dados que têm sido propostos, encontram-se os clássicos, o hierárquico, rede, relacional e orientado a objeto. Neste texto nos restringiremos a apresentar os modelos relacional e o modelo orientado a objeto.

### ***Modelo Relacional***

Modelo de dados mais popularmente conhecido e utilizado tanto nos ambientes de pesquisa das universidades quanto no processamento de dados comercial.

Segundo (Korth & Silberschtz 1993) o modelo relacional representa o banco de dados como uma coleção de tabelas, cada qual designada por um nome único. Uma linha numa tabela representa um relacionamento entre um conjunto de valores. Uma vez que uma tabela é uma coleção de tais relacionamentos, existe uma correspondência entre os conceitos de tabela e conceito matemático de relação. Na terminologia utilizada pelo modelo relacional, uma linha corresponde a uma tupla, uma coluna a atributo e uma relação corresponde a uma tabela. Alguns conceitos importantes para o entendimento o modelo relacional serão definidos a seguir.

**Tabela:** os matemáticos definem uma relação como sendo um subconjunto de um produto cartesiano de uma lista de domínios. Isto corresponde quase exatamente à definição de tabela. A diferença é que são designados nomes a atributos, ao passo que os matemáticos, contam com identificadores numéricos, usando o inteiro 1 para representar o atributo cujo o domínio aparece primeiro na linha de domínios, 2 para o atributo cujo domínio aparece em segundo lugar e assim por diante.

**Domínio:** um domínio  $D$  é o conjunto de valores atômicos ou indivisíveis. O domínio representa os tipos de valores que podem ser atribuídos às colunas de uma tabela.

### ***Modelo Orientado a Objeto***

Uma importante meta do modelo orientado a objeto é estabelecer uma correspondência direta entre o mundo real e os objetos do banco de dados de tal forma que os objetos não percam sua identidade e possam facilmente ser identificados e manipulados. A implementação interna de um objeto inclui a definição de estruturas de dados e operações para acessar essas estruturas (métodos). Alguns conceitos básicos importantes incluídos no modelo serão definidos a seguir..

- **Objetos e identidade:** cada entidade do mundo real é modelada como um objeto. Para cada objeto armazenado no banco de dados é associado um identificador único:
- **Objetos persistentes:** Objetos em linguagem de programação orientada a objeto existem apenas durante a execução do programa. O modelo objeto provê meios para criar objetos persistentes, que podem existir permanentemente armazenados em memória secundária e podem ser compartilhados por vários programas e aplicações
- **Objetos complexos:** A principal motivação para o desenvolvimento de sistemas orientados a objetos foi o desejo de representar objetos complexos em um banco de dados. Existem dois tipos de objetos complexos, objetos estruturados e não estruturados. O conceito de objeto estruturado permite que objetos complexos seja definidos através de outros objetos. O conceito de objeto complexos não estruturados é um tipo de dados que requer um grande espaço de armazenamento, como dados do tipo que representa uma imagem ou um objeto do tipo texto.
- **Encapsulamento:** termo formal que descreve a junção de métodos e dados dentro de um objeto de maneira que o acesso aos dados seja permitido apenas por meio dos próprios métodos do objeto. Quando aplicado a banco de dados garante que a estrutura interna de um objeto seja escondida e só possa ser acessada através de operações (métodos). Podemos definir operações para criar, destruir objetos, atualizar valores de objetos (ou estado do objeto), consultar parte do valor de um objeto ou ainda para aplicar cálculos nos seus valores. Este conceito não se aplica a bancos de dados relacionais no qual é predefinido um



conjunto de operações as quais podem ser aplicadas a todos os tipos de objetos. A relação e os atributos são visíveis para usuários e, programas externos acessam a relação, as tuplas e os atributos por meio dessas operações..

- **Classes e Objetos:** uma classe define um conjunto de objetos *run-time*, ou instâncias, cada um dos quais é caracterizado por operações (métodos), e o valor dos dados internos (variáveis de instâncias).
- **Herança:** uma classe pode ser definida como instância de uma ou mais classes existentes e herdar os atributos e métodos de cada classe.
- **Tipo Construtor:** conjunto de operações e estruturas básicas que podem ser combinadas para formar objetos simples ou complexos.

### *Comparação entre as representações dos modelos de dados*

Segundo (Elmasri & Navathe, 1994) a principal diferença entre os modelos de dados está relacionada com a forma como os relacionamentos são implementados. No **modelo relacional**, as conexões entre duas relações são representadas por uma chave estrangeira em uma relação que referencia a chave primária em outra relação. Tuplas individuais que têm correspondência de valores em termos de suas chaves primárias e estrangeiras estão logicamente associadas, apesar de não estarem fisicamente conectadas

Os sistemas baseados **no modelo de objetos** o relacionamento entre os **objetos** são tipicamente representados por referências a **identidades dos objetos**. Semelhantemente ao que é feito no modelo relacional quando definimos chaves estrangeiras, o modelo de objetos suporta **objetos complexos** (tuplas, conjuntos, listas e outros construtores). Em adição suporta a especificação de **métodos** e **mecanismos de herança** que permitem a criação de uma nova classe a partir de outra existente. Na Figura 4.3 é apresentado um sumário comparativo de algumas das características presentes nos modelos relacional e orientado a objeto.

	Modelos de dados	
	Relacional	Orientado a Objeto
Estrutura de dados	Esquema da relação	Descrição da classe
Instância	Tupla	Objeto
Relacionamento de 1:1	Definição de chave primária em uma relação e chave estrangeira na outra	Referência a pares de atributo inverso, ambos são simples
Relacionamento de 1:N	Definição de chave primária em uma relação e chave estrangeira nas outras	Referência a pares de atributo inverso, um deles é um conjunto válido
Relacionamento de M.N	Definição de uma relação que inclui como chaves estrangeiras as chaves primárias das relações participantes	Referência a pares de atributo inverso, onde ambos são conjuntos válidos
Conjunto de valores	Domínio	Tipo de dados atômico

Figura 4.3. Quadro comparativo dos modelos relacional e objeto

## 2.2 Linguagens de consulta

Uma das mais importantes características de um sistema de banco de dados é o provimento de uma linguagem para definição, recuperação e manipulação de dados. Os sistemas de banco de dados relacional fornecem duas linguagens separadas, uma linguagem de programação para desenvolvimento de aplicações e uma linguagem de consulta que inclui DDL, DML e sublinguagens. Os sistemas orientados a objeto provêm uma linguagem de banco de dados com características de linguagem de consulta e de programação.

Os sistemas orientados a objetos provêm uma linguagem para acesso a banco de dados que é computacionalmente completa, isto é, a linguagem de banco de dados pode executar operação que antes só podiam ser executadas pelas linguagens de programação.

O modelo relacional oferece muitas linguagens de alto nível. A partir das operações formais da álgebra relacional que podem ser aplicadas a um conjunto de tuplas; uma consulta é especificada através de uma sequência de operações sobre uma relação. No cálculo relacional utiliza-se uma única expressão para especificar uma consulta, especificando-se, apenas o que deve ser recuperado do banco de dados, não sendo especificado como fazê-lo. As linguagens comerciais (SQL, QUEL e QBE), implementadas para o modelo relacional baseiam-se no cálculo relacional. Estas linguagens oferecem também, facilidades para funções aritméticas, de agrupamento, ordenação e duplicação de tuplas. Os sistemas relacionais fornecem um conjunto de

facilidades para definição e manipulação do banco de dados em ambiente interativo ou através de comandos embutidos na linguagens de programação (COBOL, C, Pascal).

Nos sistemas orientados a objeto, a DML está embutida em linguagens de programação orientada a objeto como C++. Portanto, a forma de armazenamento de objetos transientes (definidos em programas de aplicação) e objetos persistentes é compatível. Esta característica torna-se bastante desejável quando consideramos a integração de bancos de dados com sistemas de software complexos. Linguagens de consulta têm sido desenvolvidas para sistema de banco de dados orientados a objeto. O ODMG desenvolveu uma linguagem de consulta de objetos chamada OQL, a qual suporta o modelo de dados ODMG.

### 2.2.2 Padrões de Linguagem de Consulta

Originalmente, SQL foi chamada de Sequel, foi implementada como parte do *System R* e hoje estabeleceu-se claramente como a linguagem padrão para banco de dados de dados relacional. Notações de SQL tem sido implementada por muitos fabricantes de SGBD. A junção de esforços para definição de uma versão de SQL baseado no padrão ANSI (*American National Standards Institute*) e ISO (*International Standards Organization*) tem constituiu-se o principal fator para definição da primeira versão do SQL padrão. (ANSI, 1986ab) chamada SQL1. Mais tarde esta versão foi revisada e bastante expandida, dando origem ao SQL2 (ANSI, 1989), também referenciada por SQL-92. O SQL3 é a mais nova versão do SQL, como extensão do SQL2 para incorporação dos conceitos de orientação a objeto.

O grupo ODMG-93 em (Castell, 1996) descreve a linguagem de consulta OQL (*Object Query Language*) e a linguagem de definição de objetos ODL (*Object Definition Language*). A OQL é uma linguagem de consulta para banco de dados orientado a objetos que suporta o modelo de dados definido pela ODMG. A ODL, é uma linguagem de definição de dados usada para definir tipos de objetos suportada pelo modelo de objeto ODMG. O principal objetivo desta linguagem é facilitar a portabilidade de esquemas de banco de dados entre os sistemas de bancos de dados orientados a objeto. A ODL também prover facilidades para promover a interoperabilidade entre SGBDOOs de fabricantes diferentes.

### 3. Administração do SGBD

Neste grupo estão relacionadas recursos oferecidos pelo SGBD para garantir: (1) o controle de concorrência e recuperação dos dados, (2) permitir a especificação de restrição de integridade, (3) garantir a otimização de consulta, (4) assegurar a segurança dos dados.

#### 3.1 Controle de concorrência e recuperação

Segundo (Cattell, 1994) mecanismos de controle de concorrência e recuperação são importantes características de um SGBD. Estes mecanismos facilitam, respectivamente a manutenção lógica e física do banco de dados.

Mecanismos de recuperação permitem que em caso de queda do sistema (pique de energia, defeitos físicos no *hardware*), possa ser integrado o estado consistente do banco de dados. Porque, os dados em um banco de dados têm vida longa, se comprado com outras estruturas de dados que normalmente residem na memória apenas por alguns segundos. Os *backups* são os mais simples e antigos mecanismos de recuperação. Estas cópias de segurança normalmente chamadas de *backup* são geradas em intervalos de tempo regulares. Por exemplo, todo final de tarde ou noite, possibilitando que em caso de falha irreparável do sistema não mais que um dia de trabalho seja perdido, através da recuperação da cópia de segurança, ou ainda através daquilo que chamamos de *backups* incrementais onde pode ser recuperada apenas aquela parte do banco de dados que foi danificada quando da falha do sistema.

Um dos conceitos mais importantes em sistemas modernos é a multiprogramação. Tendo em vista que, diversas transações podem ser executadas concorrentes e o processador pode ser compartilhado entre elas. Assim para garantir a sincronização e o isolamento de transações concorrentes, assegurando a consistência do banco de dados, na execução dessas transações, são implementadas nos sistemas gerenciadores de banco de dados técnicas baseadas em protocolos ou conjunto de regras. Aqui comenta-se algumas dessas técnicas.

As técnicas mais utilizadas pelos sistemas relacionais para implementação de controle de concorrência baseiam-se nos mecanismos de bloqueio e marcadores de tempo, serão descritos a seguir.

**Protocolos baseados em bloqueio Loking** . Uma das principais técnicas usadas para sincronização na execução de transações concorrentes em um banco de dados, se baseia no bloqueio de itens de dados. Um protocolo de baseado em bloqueio é um conjunto de regras que estabelece quando uma transação pode bloquear e desbloquear cada um dos itens do banco de dados.

- Bloqueio Binário. Se uma transação  $T_i$  necessita realizar uma operação de leitura ou escrita em um item do banco de dados, será associado o valor 1 a variável de bloqueio deste item, então se uma transação  $T_j$  requisitar um bloqueio neste mesmo item, esta ficará no estado de espera, até que  $T_i$  libere o item, atualizando a variável de bloqueio para zero. Por outro lado  $T_i$  pode requisitar um bloqueio em um item de dados que tenha sido bloqueado por  $T_j$ , tendo  $T_i$  que ficar no estado de espera até que o item de dados seja liberado.
- Bloqueio partilhado. Se uma transação  $T_i$  obteve um bloqueio no modo partilhado, no item  $Q$ , então  $T_i$  pode ler este item, mas não pode gravar  $Q$ , pois esta operação só será permitida quando  $T_i$  obtiver um bloqueio exclusivo para  $Q$ .
- Bloqueio exclusivo Se uma transação  $T_i$  obteve um bloqueio no modo exclusivo no item  $Q$ , então  $T_i$  pode ler e gravar  $Q$ ,
- Bloqueio em duas fases. Este protocolo requer que bloqueios e desbloqueios de itens de dados devam acontecer em duas fases: crescimento e encolhimento. Inicialmente a transação está na fase de crescimento, a transação adquire então tantos bloqueios quantos forem necessários. Uma vez que a transação libere um bloqueio, ela entra na fase de encolhimento e não pode requerer bloqueios de itens de dados.

O principal problema identificado nos protocolos de bloqueio apresentado é o impasse ocorrido entre duas transações conflitantes. Isto é, se uma transação  $T_i$  obtém um bloqueio no modo exclusivo no item  $Q$  e uma transação  $T_j$  está requisitando um bloqueio no modo partilhado em  $Q$ , esta terá que ficar em estado de espera até que  $T_i$  possa liberar  $Q$ . Por outro lado, se  $T_i$  está requisitando um bloqueio no modo partilhado para o item de dados  $P$  o qual já havia sido bloqueado no modo exclusivo por  $T_j$  então  $T_i$  terá que ficar no estado de espera até que  $T_j$  libere  $P$ , causando assim uma situação em que as duas transações terão seu processamento interrompido. Esta situação pode ser resolvida se uma das transações for reprocessada. Uma das soluções utilizadas pelos SGBDs é a

implementação de protocolos baseados em marcadores de tempo, os quais nunca pedem que uma transação entre num estado de espera.

- **Protocolos baseados marcadores de tempo *Timestamp*.** Um marcador de tempo pode ser implementado de suas formas, utilizando o relógio do sistema, onde o marcador equivale ao valor marcado pelo relógio quando a transação entrar no sistema, ou pode ser implementado utilizando contador lógico que seja incrementado depois que um novo marcador de tempo for designado, isto é um marcador tempo neste caso é equivalente ao valor do contador quando a transação entrar no sistema. Protocolos de ordenação de marcadores de tempo asseguram que, qualquer operação de leitura ou escrita conflitantes, em itens de dados, seja executada na ordem do marcador de tempo. Assim, protocolos baseados em marcadores de tempo evitam impasses, uma vez que as transações nunca são forçadas a entrar no estado de espera., por muito tempo.

Os sistemas gerenciadores de banco de dados orientados a objetos oferecem novos mecanismos para garantir o controle de concorrência e recuperação de banco de dados, porque os SGBDs foram concebidos com o objetivo de atender a usuários com novos tipos de requerimentos. Por exemplo, várias versões de um mesmo objeto podem ser armazenadas em um banco de dados, e o conhecimento semântico dos objetos pode influenciar bastante na granularidade desejável e no tipo de controle de concorrência

### 3.2 Restrições de Integridade

**Restrições de integridade** asseguram que mudanças feitas no banco de dados por usuários autorizados não resultem em inconsistência de dados. Aqui comentam-se os tipos de restrições que podem ser implementadas em SGBDs baseados nos modelos: relacional e orientado a objeto.

Aqui nós discutiremos alguns tipos de restrições que podem ser especificadas no esquema de banco de dados baseado no **modelo relacional** através de comandos oferecidos pela linguagem de definição de dados do SGBD ou, através da especificação de regras. Estes incluem restrições de domínio, integridade de entidade e restrições de integridade referencial.

- **Restrições de domínio:** garantem que os valores de cada atributo  $A$  devem ser um valor atômico de um domínio  $dom(A)$  para aqueles atributos, que devem ser suportados pela Linguagem *SQL*, utilizada pelos SGBDs relacionais. Os tipos de

domínio suportado pela linguagem *SQL* incluem o tipos: inteiros (*short-integer*, *integer*, *long-integer*), real (*float* e *double float*), caracter, cadeia de caracter de tamanho fixo, cadeia de caracter de tamanho variável, data, hora, *timestamp*.

- **Restrições de entidade.** Garante que nenhuma chave primária pode assumir um valor nulo, porque o valor da chave primária é usado para identificar uma única tupla em uma relação.
- **Restrições de integridade referencial.** Este tipo de restrição é usado para garantir a consistência entre duas relações. Por exemplo, imagine o banco de dados de recursos humanos de uma empresa, composto pelas relação **Empregado** com os atributos: **matrícula**, **identificação do cargo** e **salário** e a relação **Cargo** com os atributos: **identificação do cargo** e **descrição**. Na relação **Cargo** o atributo *identificação do cargo* foi definido como chave primária, enquanto que na relação **Empregado** o atributo **identificação do cargo** foi definido como chave estrangeira. Assim, cada instância do atributo **identificação do cargo** na relação **Empregado** deve corresponder a uma instância do atributo **identificação do cargo** na relação **Cargo**.

Os tipos de restrições suportados pelos SGBDs orientados a objetos variam de sistema para sistema. Filosoficamente as restrições seriam especificadas nos métodos, os quais são bastante genéricos, dificultando a verificação de quais restrições foram especificadas. O Mecanismo de relacionamento inverso suportado por muitos sistemas orientados a objetos permite a especificação daquilo que chamamos de restrição de relacionamento.

Um importante conceito utilizado na definição de restrições de integridade no modelo objeto é o de relacionamento inverso. Analogamente ao que é feito no modelo relacional onde são definidos atributos chaves para representação de relacionamento entre atributos de uma relação, no modelo de objeto são definidos novos atributos para representar relacionamento entre objetos. Um relacionamento inverso é simplesmente a definição de um novo atributo em um objeto que irá fazer referência a um atributo do outro objeto, indicando que existe um relacionamento entre esses objetos.

Por exemplo, o relacionamento existente entre os objetos documentos e capítulos de um livro, pode ser representado pelo atributo **doc** do objeto capítulo e **cap** do objeto

documento (ver Figura 4.4), indicando que cada capítulo pode ter vários documentos, mas cada documento só pode pertencer a um capítulo. Este tipo de relacionamento como será visto mais adiante pode ser utilizado como forma de definir restrições de integridade referencial.

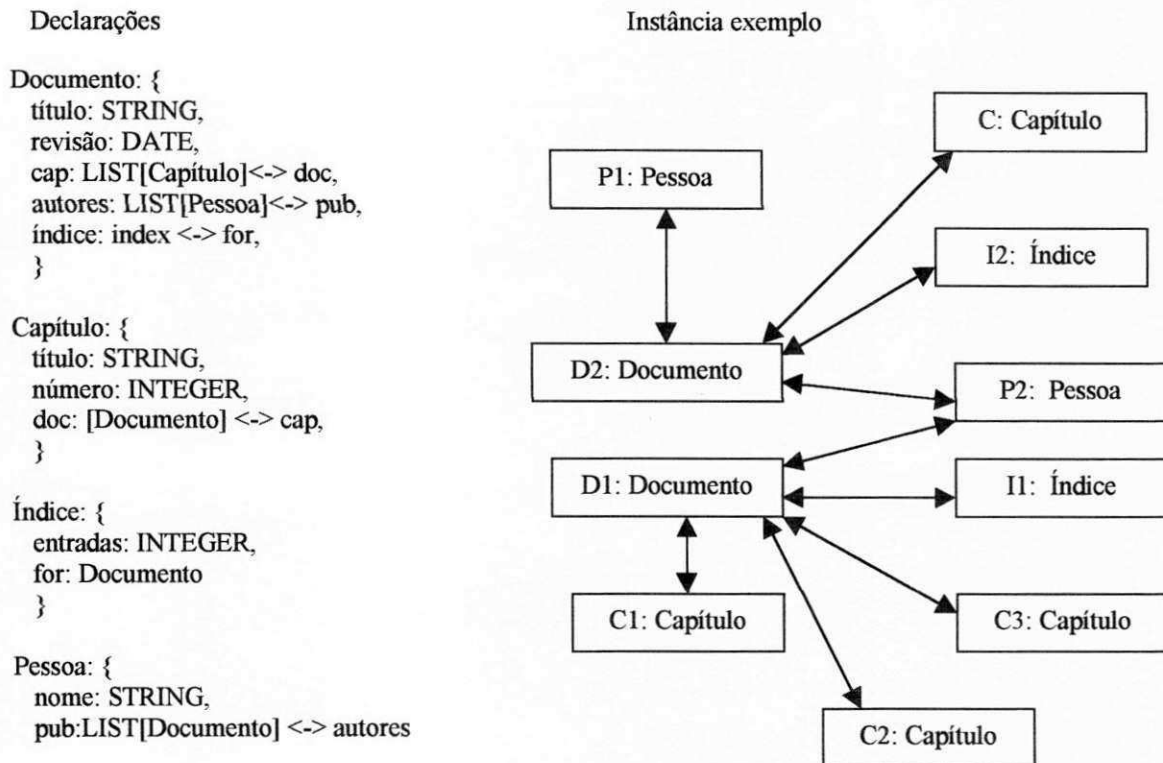


Figura 4.4 - Representação de relacionamento em um modelo de dados orientado a objeto.

### 3.3 Otimização de consulta

**Otimização de consulta** resulta de um conjunto de recursos que permite ao SGBD criar uma estratégia de execução para recuperação do resultado de uma consulta, expressa em uma linguagem de consulta de alto nível como o SQL, nos arquivos internos do banco de dados. É alcançada através da escolha de uma estratégia de execução apropriada, dentre as diversas estratégias possíveis, para o processamento de uma consulta (Elmasri & Navathe, 1994). As principais técnicas utilizadas pelos desenvolvedores na implementação das técnicas para otimização de consulta usualmente se baseiam na combinação de duas estratégias: a regra heurística para ordenação das operações em uma estratégia de execução de consulta e a estimativa de custo (Elmarri, 1994)



- **regra heurística** reordena as operações de uma árvore de consulta ou determina a ordem para execução especificada no grafo de consulta ;
- **estimativa de custo** estimativa sistemática dos custos de diferentes estratégias de execução e escolha do plano de execução com a menor estimativa de custo.

### 3.4 Segurança

Os sistemas de banco de dados multiusuários devem prover mecanismos para proteger os dados armazenados no banco de dados para garantir que certos usuários ou grupo de usuários acessem porções selecionadas de um banco de dados. Isto é, particularmente importante, por exemplo, quando se tem em uma corporação um grande volume de dados armazenados em um banco de dados e várias categorias de usuários com permissão para acessar determinadas porções destes dados. Os SGBDs normalmente provêm dois tipos mecanismos para garantir esta segurança.

- **Mecanismo de segurança discreto:** usado para conceder privilégios aos usuários do banco de dados, incluindo acesso a arquivos de dados específicos, registros, ou arquivos nos modos de leitura, escrita ou atualização.
- **Mecanismo mandatório de segurança:** usado para definir vários níveis de segurança através da classificação de dados e de usuários dentro de várias classes (ou níveis) de segurança utilizadas na implementação da política de segurança da organização

Um outro problema de segurança muito comum em qualquer sistema de computação é a prevenção de acessos ao sistema de pessoas não autorizadas para obtenção de informações ou realizar mudanças maliciosas em porções do banco de dados. Os mecanismos de segurança do SGBD devem incluir previsões de restrição de acesso ao sistema de banco de dados como um todo. Esta função é chamada de **controle de acesso**, manuseada através da criação de uma “conta” (onde serão informadas identificação e senha) e através da qual o usuário será identificado pelo sistema, no momento que tentar acessá-lo.

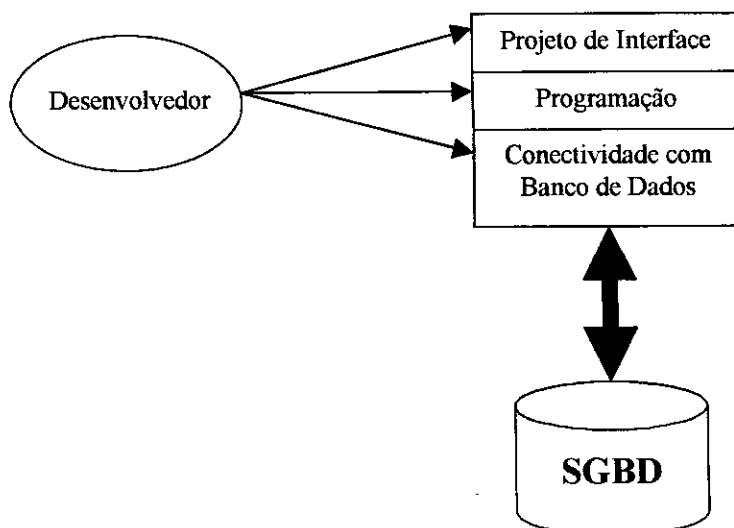
O mecanismo **criptografia**, a qual é usada para proteger dados suscetíveis a acessos inescrupulosos que são normalmente transmitidos via satélite ou qualquer outro

tipo de comunicação de rede. Criptografia pode ser usada para prover proteção adicional a porções do banco de dados. Os dados são codificados usando algum tipo de algoritmo. Um usuário não autorizado que tente acessar dados codificados desta maneira terão dificuldades para decifrá-los, mas usuários autorizados terá acessos aos algoritmos para decodificação (ou chaves) para decifrá-los.

#### 4. Ambiente de desenvolvimento

A este grupo estão associados componentes e características básicas de um ambiente rápido de desenvolvimento de banco de dados (RAD – *Rapid application development*) e integrado (IDE – *integrated development environment*). Um ambiente IDE permite aos projetistas de interface construir as interfaces, desenvolver programas e definir minuciosos componentes de programação, através de recursos como, linguagem de programação visual orientada a eventos, depurador de código, editor de imagem, gerador de relatório, gerador de formulário, linguagem de programação procedural (Pascal, COBOL,) ou orientada a objetos (C++, Java).

Com objetivo de auxiliar o *decion maker* no entendimento do complexo mundo das ferramentas de desenvolvimento de aplicação para banco de dados, relaciona-se um conjunto de características que podem ser encontradas na maioria dessas ferramentas, apresentadas aqui, em três camadas ou subsistemas básicos (Figura 4.5): camada de conectividade com banco de dados, camada de projeto de interface e a camada de programação.



**Figura 4.5 - Camadas básicas de ambientes de desenvolvimento para banco de dados**

#### 4.1. Conectividade com banco de dados

Segundo Linthicum e Shumacher em (Linthicum, 1996, 1997) e (Shumacher, 1996), a camada de **conectividade com banco de dados** atua como uma interface entre o banco de dados e as outras camadas. A camada de conectividade suporta todas as chamadas ( comandos de DDL e DML) de baixo nível do banco de dados para a ferramenta. Esta camada pode tornar a aplicação independente do banco de dados, sendo necessário apenas que o ambiente de desenvolvimento seja capaz de carregar e descarregar os *drivers* requeridos para acessar o banco de dados.

A camada de conectividade permite que a ferramenta acesse os dados através de aplicações nativas orientadas a objetos, manipulando-se objetos (mudando propriedades e invocando métodos), através de esquemas relacionais, manipulando tabelas, colunas, linhas e relacionamentos ao invés de trabalhar com objetos, e finalmente, usando *drivers* nativos ou simplesmente ignorando os (*middleware* como ODBC e JDBC e indo diretamente para a API nativa do banco de dados).

#### 4.2. Projeto de Interface

A camada de **projeto de interface** é o subsistema responsável por criar as telas, janelas e menus através dos quais o usuário pode interagir com a aplicação.

#### 4.3. Programação

A camada de **programação** da ferramenta permite aos engenheiros de sistemas customizar os componentes das aplicações utilizando recursos de programação visual orientada a eventos ou geração de código através de linguagens de programação procedural (Pascal, COBOL,) ou orientada a objetos (C++, Java).

A camada de programação permite que as aplicações geradas possam ser portáteis para várias máquinas e sistemas operacionais incluindo Windows, UNIX, *Macintosh* OS, O suporte a padrões é um importante recurso que pode ser oferecido pelo subsistema de programação

A Figura 4.6 ilustra um exemplo de refinamento da árvore de requisitos para o nó 2.3 do grupo Ambiente de Desenvolvimento da árvore AGR.

2. Ambiente de Desenvolvimento	2.3. Linguagem de programação	2.3.1. Pascal
		2.3.2 C++
		2.3.3 COBOL

Figura 4.6 Exemplo de refinamento da AGR

#### 4.1.3.2. Elaboração do Questionário

Nesta fase, sugere-se a elaboração de um questionário destinado ao fornecedor, que contemple os requisitos presentes na árvore AER. Este questionário deverá incluir questões subjetivas e objetivas necessárias a avaliação do fornecedor e do SGBD. No Apêndice II encontra-se um exemplo deste tipo de questionário utilizado no estudo de caso descrito no Capítulo 5.

#### 4.1.3.3. Classificação dos Grupos de Requisitos

Nesta fase é feita a classificação dos grupos de requisitos da árvore AER em: essenciais, desejáveis e opcionais a exemplo da classificação utilizada no modelo LSP (Su et al. 1987). As Figuras 4.7 (a) e 4.7 (b) exemplificam esta classificação.

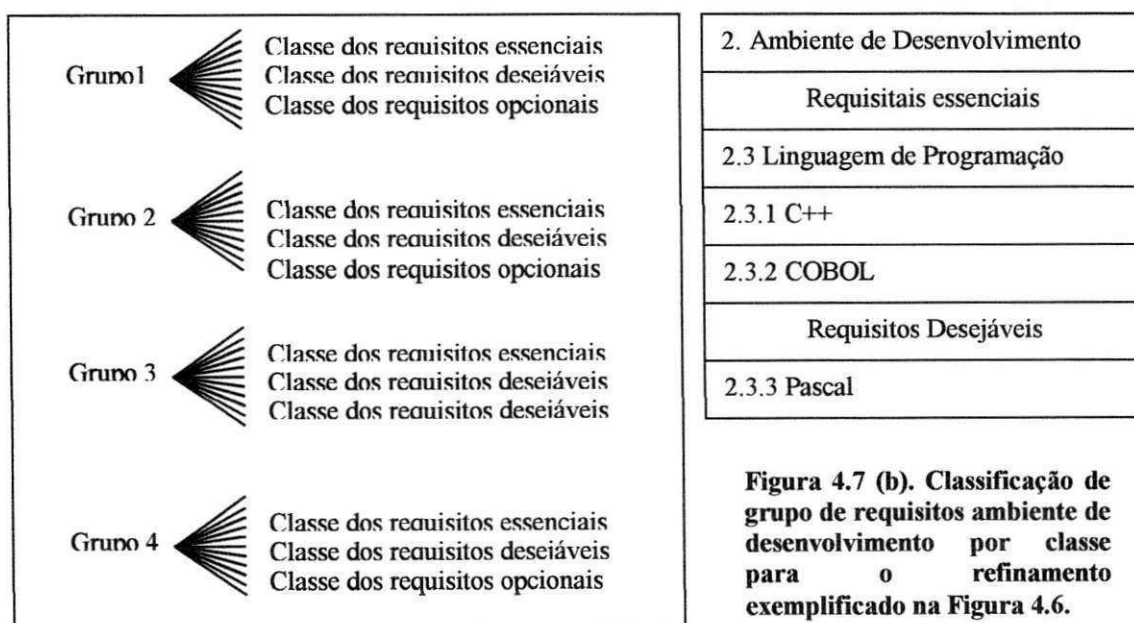


Figura 4.7 (b). Classificação de grupo de requisitos ambiente de desenvolvimento por classe para o refinamento exemplificado na Figura 4.6.

Figura 4.7 (a). Representação Gráfica da Classificação dos Grupos de Requisitos em Classes de Requisitos Essenciais, Opcionais e Desejáveis

- **Classe dos requisitos essenciais** diz respeito àquelas características que são indispensáveis a um SGBD para que este possa atender às necessidades da organização. Por exemplo, uma organização que possui um ambiente de banco de dados cliente/servidor necessita que o produto a ser adquirido atenda a esta especificação.
- **Classe dos requisitos desejáveis** são requisitos extremamente desejáveis, mas não absolutamente essenciais para que o produto seja escolhido e portanto de menor prioridade para determinar esta escolha.
- **Classe dos requisitos opcionais** referenciam requisitos cuja a presença ou ausência não afeta significativamente a nota final atribuída ao produto na etapa de classificação dos produtos.

#### **4.1.4. Classificação dos Produtos**

A classificação dos produtos baseada nos requisitos presentes no questionário tem como objetivo indicar quais produtos são capazes de atender às necessidades do usuário. Esta fase terá como saída uma lista de produtos. A lista é construída a partir dos resultados das análises que se baseiam nas informações do fornecedor, nos testes de desempenho e nos custos. A lista de produtos será utilizada na próxima fase, pelo *decision-maker*, para que, com base nos critérios de desempate, seja possível escolher o produto mais adequado.

##### **4.1.4.1. Classificação Baseada nas Informações do Fornecedor**

A classificação com base nas informações do fornecedor ocorre em duas etapas: (1) análise dos requisitos quantificáveis, e (2) análise dos requisitos não quantificáveis.

##### **Classificação dos Grupos de Requisitos Quantificáveis**

Esta análise tem como finalidade a classificação dos SGBDs alternativos a partir da atribuição de “notas”, em função do atendimento ou não aos grupos de requisitos presentes na AER.

Para sintetizar o processo de análise dos requisitos quantificáveis este procedimento foi dividido em três estágios:

Agregação dos requisitos por grupo;

1. Agregação de um grupo de requisitos por classe;
2. Agregação de uma classe de requisitos;
3. Cálculo da “nota” a ser atribuída a cada SGBD.

### Estágio 1. Agregação de um grupo de requisitos por classe

O primeiro estágio tem como objetivo realizar a agregação dos grupo de requisitos: aspectos gerais, ambiente de desenvolvimento, administração dos dados e administração do SGBD, pertencentes a uma mesma classe;

- a. Agregação de um grupo da classe de requisitos essenciais: a agregação de uma classe de requisitos essenciais é feita com base no produtório, representado na expressão 4.1, onde  $n$  representa o número de requisitos (ao nível das folhas) na classe dos requisitos essenciais,

$$E_x = \prod_i^n e_i \quad (4.1)$$

onde  $x$  representa cada um dos grupos de requisitos: *AG* Aspectos Gerais, *AD* Ambiente de Desenvolvimento, *ADD* Administração dos Dados e *AS* Administração do SGBD.

Na expressão 4.1,  $e_i$  representa o  $i$ -ésimo requisito essencial de um grupo, podendo assumir os valores 0 ou 1. A variável  $e_i$  assumirá o valor 0 (zero) quando o requisito não for oferecido pelo SGBD e, assumirá o valor 1 (um) quando o requisito for oferecido pelo SGBD.

- b. Agregação de um grupo da classe de requisitos desejáveis a agregação de uma classe de requisitos desejáveis é feita com base na média aritmética, representada na expressão 4.2, onde  $n$  representa o número de requisitos (ao nível das folhas) na classe dos requisitos desejáveis, e  $x$  representa cada um dos grupos de requisitos.

$$D_x = \frac{\sum_i^n d_i}{n} \quad (4.2)$$

Nesta expressão, que é análoga à expressão 4.1,  $d_i$  representa o  $i$ -ésimo requisito do grupo, podendo assumir os valores 0 ou 1,  $d_i$  assumirá o valor 0 (zero) quando o requisito não for satisfeito pelo SGBD e assumirá o valor 1 (um) quando o requisito for oferecido pelo SGBD.

- c. Agregação de um grupo da classe de requisitos opcionais a agregação de uma classe de requisitos opcionais é feita com base na média aritmética, representada na expressão 4.3, onde  $n$  representa o número de requisitos (ao nível das folhas) na classe dos requisitos opcionais e  $x$  representa cada um dos grupos de requisitos

$$O_x = \frac{\sum_1^n o_i}{n} \quad (4.3)$$

Por analogia,  $o_i$  representa o  $i$ -ésimo requisito opcional do grupo, podendo assumir os valores 0 ou 1. A variável  $o_i$  assumirá o valor 0 quando o requisito não for satisfeito pelo SGBD e assumirá o valor 1 quando o requisito for satisfeito pelo SGBD.

### **Estágio 2. Agregação de uma classe de requisitos**

Este estágio da metodologia tem como objetivo a agregação dos grupos de requisitos pertencentes à mesma classe

- a. Agregação das classes de requisitos essenciais é feita com base na expressão 4.4:

$$A_{RE} = \frac{\sum_1^m E_x}{m} \quad (4.4)$$

onde  $A_{RE}$  é a agregação das classes de requisitos essenciais,  $E_x$  (expressão 4.1) é o valor referente a cada classe de requisitos essenciais e,  $m$  é o número de classes de requisitos.

b. Agregação das classes de requisitos desejáveis é feita com base na expressão 4.5:

$$A_{RD} = \frac{\sum_1^m D_x}{m} \quad (4.5)$$

onde  $A_{RD}$  é a agregação das classes de requisitos desejáveis,  $D_x$  (expressão 4.2) é o valor referente a cada grupo de requisito desejável e  $m$  é o número de grupos de requisitos.

c. Agregação das classes de requisitos opcionais é feita com base na expressão 4.6:

$$A_{RO} = \frac{\sum_1^m O_x}{m} \quad (4.6)$$

onde  $A_{RO}$  é a agregação das classes de requisitos opcionais,  $O_x$  (expressão 4.3) é o valor referente a cada grupo de requisito opcional e  $m$  é o número de grupos de requisitos.

### Estágio 3. Cálculo da “nota” a ser atribuída a cada SGBD

Este estágio tem como objetivo o cálculo da nota final atribuída a cada SGBD, através da expressão 4.7. Para garantir que a nota final expresse a importância relativa à cada classe de requisito, foram atribuídos pesos diferenciados para cada classe.

$$N_{FINAL} = 6A_{RE} + 3A_{RD} + A_{RO} \quad (4.7)$$

onde,  $A_{RE}$  (expressão 4.4) representa a agregação dos grupos de requisitos essenciais,  $A_{RD}$  (expressão 4.5) representa a agregação de requisitos desejáveis e  $A_{RO}$  (expressão 4.6) representa a agregação de requisitos opcionais.



### **Classificação dos Requisitos Não Quantificáveis**

Esta análise tem como propósito a classificação dos SGBDs, no que se refere ao atendimento aos requisitos não quantificáveis. Ver Tabelas 4.2 (a) e 4.2 (b), contendo a lista de produtos associados aos índices de satisfação (ver Tabela 4.1). Estes índices representam o nível de satisfação do *decision-maker* face às respostas dos fornecedores para os quesitos não quantificáveis do questionário Apêndice II.

Índice de Satisfação	Descrição
1	Resposta Satisfatória
2	Resposta Insatisfatória
3	Ausência de Resposta

**Tabela 4.1 - Lista de índices de satisfação**

### **Agregação dos Requisitos Não Quantificáveis**

Para facilitar, a análise e classificação dos requisitos não quantificáveis foi dividida em dois estágios:

1. Agregação dos requisitos não quantificáveis
2. Classificação dos produtos baseada na análise dos requisitos não quantificáveis

#### **Estágio 1 - Agregação dos requisitos não quantificáveis**

Os requisitos não quantificáveis são agregados com base na expressão 4.8, onde  $r$  representa o valor do  $i$ -ésimo requisito não quantificável e  $n$  é o número de requisitos não quantificáveis analisados (ver Tabela 4.2 (a)).

$$T_{Pontos} = \sum_{i=1}^n r_i \quad (4.8)$$

## Estágio 2 - Classificação dos produtos baseada na análise dos requisitos não quantificáveis

Este estágio consiste na classificação dos produtos. Alcançará o primeiro lugar aquele produto que apresentar o menor valor numérico como resultado da agregação, (ver Tabela 4.2 (b)).

Lista de Quesitos Não Quantificáveis	Lista de Produtos associados aos índices de satisfação		
	SGBD 1	SGBD 2	SGBD 3
Última versão no Brasil e no Exterior	1	2	2
Recursos Mínimos de Hardware e Software	2	1	2
Relação de Cursos	3	1	1
Total de Pontos	6	4	5

Tabela 4.2 (a) - Lista de produtos baseada no atendimento aos requisitos não quantificáveis

Classificação dos Produtos Baseada na Análise de Requisitos Não quantificáveis	
Nome do Produto	Classificação
SGBD 2	Primeiro Lugar
SGBD 3	Segundo Lugar
SGBD 1	Terceiro Lugar

Tabela 4.2 (b) - Exemplo de uma tabela de produtos classificados com base na Análise de requisitos não quantificáveis

### 4.1.4.2. Classificação Baseada nos Testes de Desempenho

Esta fase objetiva a realização de uma análise de desempenho dos SGBDs classificados nas fases anteriores, através da aplicação de técnicas de *benchmark*. Ao final desta fase teremos como saída uma lista de SGBDs, ordenados de acordo o seu desempenho, como exemplificado na Tabela 4.3.

Classificação dos Produtos Baseada na Análise de Desempenho	
Nome do Produto	Classificação
SGBD 1	Primeiro Lugar
SGBD 2	Segundo Lugar
SGBD 3	Terceiro Lugar

Tabela 4.3 - Exemplo de uma tabela de produtos classificados com base na análise de desempenho

#### **4.1.4.2.1 Análise de Desempenho Baseada na Metodologia de *Benchmark* Especializado**

Dentre as metodologias de *benchmark* de banco de dados mencionadas no Capítulo 3, consideramos a Metodologia de *Benchmark* Especializado (MBE), (Pereira, 1990) como a mais indicada para o usuário da MASS. O motivo da escolha é sua flexibilidade, permitindo que o usuário parametrize o *benchmark* de acordo com as características de uma aplicação ou de um conjunto típico de aplicações.

#### **4.1.4.2.2. Análise de Desempenho Baseada em Testes de Tomada de Tempo**

Uma segunda alternativa para teste de desempenho dos produtos, porém com resultados não tão precisos, é a realização de testes de tomada de tempo. Estes testes se baseiam na especificação de uma aplicação exemplo, a ser modelada pelo usuário da MAPESGBD, que deve ser implementada pelos fornecedores dos produtos. Por serem testes simples, têm a vantagem de serem executados com facilidade no ambiente do usuário. Estes testes têm como objetivo principal verificar o tempo de resposta para realização de operações de consulta e atualização às bases de dados.

É importante lembrar que o bom desempenho e a confiabilidade de um produto, não dependem apenas da qualidade do software, mas também dos recursos de hardware. No caso dos gerenciadores de banco de dados, o desempenho e confiabilidade podem ser bastante comprometidos se o disco rígido onde estão armazenados os dados não for adequado. O gerenciador ao realizar operações do tipo consulta ou atualização, necessita acessar o disco. Nesta ocasião, o seu desempenho pode ser comprometido significativamente, em função de um baixo desempenho do disco.

Atualmente os discos com tecnologia RAID são os mais adequados para este fim por oferecerem um bom desempenho em termos de operações de entrada e saída e um bom índice de confiabilidade em casos de falhas de hardware (Ganger et al. 1994).

#### 4.1.4.3. Classificação Baseada nos Custos

Esta etapa tem como objetivo a realização de uma estimativa de custos para aqueles SGBDs que apresentaram os melhores níveis de desempenho durante a fase de classificação dos produtos. (Si et al. 1987) apresenta algumas alternativas para calcular custos de sistemas, salientando que não foi possível encontrar na literatura um método que possa ser considerado eficaz na realização desta tarefa.

Dos métodos por ele apresentados, escolhemos aquele que sugere a construção de um vetor de custos, associando a cada requisito quantificável o seu valor em uma moeda, como exemplificado na Tabela 4.4.

Requisito	Classificação dos Produtos Baseada nos Custos		
	SGBD1	SGBD2	SGBD3
SGBD (Módulos que compõe o produto)			
Ferramentas Adicionais			
Recursos mínimos de hardware			
Recursos mínimos de software			
Suporte técnico			
Manutenção			
Treinamento			
Documentação Adicional			
Recursos Humanos			
<b>Custo Total</b>			

**Tabela 4.4 - Classificação dos produtos baseada nos custos**

#### 4.1.5 Classificação Baseada na relação Custo Benefício

Esta fase tem como objetivo a classificação final dos produtos com base na análise custo benefício. Tomados os resultados das etapas anteriores considera-se, a partir das necessidades e prioridades inerentes a cada organização, o produto que melhor atende às suas expectativas, por um custo condizente aos benefícios.

## ***Validação da Metodologia Proposta***

A validação da Metodologia proposta baseou-se em um estudo de caso realizado em uma das nove empresas participantes da pesquisa, citada no Capítulo 2. Esta Empresa demonstrou interesse em utilizar a metodologia, por encontrar-se na fase de estudos para aquisição de um novo SGBD.

A empresa, é uma empresa de grande porte do ramo de tubos e conexões, que utiliza um gerenciador de banco de dados no sistema de faturamento da empresa, em um ambiente multiusuário, com sistema operacional UNIX.

### **5.1. Estudo de Caso**

Nesta seção apresentaremos os resultados obtidos com a aplicação da metodologia na empresa piloto.

#### **5.1.1. Escolha do *Decision-Maker***

O grupo escolhido para realização do processo de seleção consistiu dos seguintes profissionais: gerente de informática, programador de aplicação e o responsável pelo setor de faturamento. O gerente de informática e o programador foram responsáveis pela aplicação da metodologia no que diz respeito aos aspectos técnicos e, o responsável pelo setor de faturamento representou a principal aplicação da empresa na fase de levantamento dos requisitos.

### 5.1.2. Pré-Seleção de SGBDs

Os SGBDs pré-selecionados para participar do processo de seleção foram: Oracle, Sybase, Informix, Progress e Ingres, por apresentarem uma grande penetração no mercado e por se adequarem a ambiente multiusuário, com sistema operacional UNIX.

### 5.1.3. Levantamento e Classificação dos Requisitos

Para realização do levantamento e classificação dos requisitos do ambiente de banco de dados e da aplicação de faturamento, foram feitas reuniões com o grupo responsável pelo processo.

#### 5.1.3.1. Refinamento da Árvore Genérica de Requisitos

Após o levantamento dos requisitos, o grupo passou a estudar a Árvore Genérica de Requisitos (AGR) acrescentando e retirando nós folhas, construindo assim a árvore Específica de Requisitos (AER) da empresa, a qual é ilustrada na Figura 5.1.

---

### AER - Árvore Específica de Requisitos

---

#### 1. Aspectos Gerais

- 1.1. Existência de grupos de usuários
- 1.2. Documentação disponível
  - 1.2.1. Guia do administrador de banco de dados
  - 1.2.2. Guia de instalação e operação
  - 1.2.3. Guia do usuário para gerenciamento de objeto
  - 1.2.4. Guia de *release* & conversões
  - 1.2.5. Manual de referência do gerador de aplicação
  - 1.2.6. Manual de referência do gerador de formulários
  - 1.2.7. Guia de referência de linguagem de programação
  - 1.2.8. Guia do usuário para o editor gráfico
  - 1.2.9. Guia para o administrador da rede
  - 1.2.10. Sumário de comandos da linguagem de consulta
  - 1.2.11. Manual de referência da linguagem de consulta
- 1.3. Treinamento
  - 1.3.1. Periodicidade dos treinamentos
    - 1.3.1.1. Mensal
    - 1.3.1.2. Trimestral
    - 1.3.1.3. Semestral
    - 1.3.1.4. Anual
- 1.4. Suporte técnico
  - 1.4.1. Período de atendimento
    - 1.4.1.1. 24 horas apenas em dias úteis
    - 1.4.1.2. 24 horas em dias úteis, feriados e finais de semana
    - 1.4.1.3. Apenas em horário comercial, em dias úteis

- 1.4.2. Formas de atendimento
  - 1.4.2.1. Telefone
  - 1.4.2.2. Internet
  - 1.4.2.3. Visita do técnico à instalação do usuário

## **2. Ambiente de Desenvolvimento**

- 2.1. Linguagem de consulta
  - 2.1.1. Integração com o dicionário de dados
  - 2.1.2. Suporte a multimídia
  - 2.1.3. Facilidade de uso (aspectos ergonômicos)
  - 2.1.4. Disponibilidade de tutorial *on-line*
  - 2.1.5. Suportar SQL embutido
- 2.2. Ferramentas disponíveis
  - 2.2.1. Consulta ad-hoc
  - 2.2.2. Gerador de aplicação
  - 2.2.3. Gerador de telas
  - 2.2.4. Gerador de relatórios
  - 2.2.5. Gerador de menus
  - 2.2.6. Gerador de gráficos
  - 2.2.7. CASE
  - 2.2.8. Manutenção do dicionário de dados
- 2.3. Linguagem de programação
  - 2.3.1. C
  - 2.3.2. C ++
  - 2.3.3. COBOL
  - 2.3.4. Fortran
  - 2.3.5. Pascal
- 2.4. Independência entre dados e programa
- 2.5. Portabilidade de código
  - 2.5.1. Windows
  - 2.5.2. OS/2
  - 2.5.3. VM
  - 2.5.4. MVS
  - 2.5.5. UNIX

## **3. Administração dos Dados**

- 3.1. Dicionário de dados
  - 3.1.1. Inclusão de regras de validação
  - 3.1.2. Consistência automática de campos
  - 3.1.3. Referência cruzada a objeto
  - 3.1.4. Geração automática de objetos
  - 3.1.5. Mecanismos de distribuição do dicionário de dados
- 3.2. Métodos de acesso aos dados
  - 3.2.1. Hash
  - 3.2.2. B-tree
- 3.3. Modelo de dados
  - 3.3.1. Relacional
  - 3.3.2. Orientado a objeto
- 3.4. Estruturas de informação relacionais
  - 3.4.1. Tabelas
  - 3.4.2. Colunas
  - 3.4.3. Índices
  - 3.4.4. Visões
  - 3.4.5. Domínios
- 3.5. Operações sobre objetos
  - 3.5.1. *Queries*
  - 3.5.2. *Joins*
  - 3.5.3. Projeções
  - 3.5.4. Uniões

- 3.5.5 Interseções
  - 3.5.6 Diferenças
  - 3.6. Integridade de dados
    - 3.6.1 Entidade
    - 3.6.2 Referencial
    - 3.6.3 Domínio
    - 3.6.4 Suporta *triggers* para outros tipos de integridade
  - 3.7. Tipos de dados suportado pela linguagem de definição de dados:
    - 3.7.1 Inteiro de dois bytes
    - 3.7.2 Inteiro de quatro bytes
    - 3.7.3 Inteiro de oito bytes
    - 3.7.4 Real de ponto flutuante
    - 3.7.5 Abstrato tipo data
  - 3.7.6 Abstrato tipo *Money*
  - 3.8. Funções implementadas pela linguagem de manipulação de dados
    - 3.8.1. Matemática
    - 3.8.2. Estatística
    - 3.8.3. Manipulação de strings
  - 3.9. Otimização de consulta
- 4. Administração do SGBD**
- 4.1. Controle de Concorrência
  - 4.2. Recursos para assegurar a integridade física dos dados
    - 4.2.1. Mecanismos para recuperação da base de dados
    - 4.2.2. Mecanismos de tolerância a falhas
    - 4.2.3. Espalhamento de arquivos
  - 4.3. Recursos para controle de acesso
    - 4.3.1. Recursos para definição de níveis de acesso
    - 4.3.2. Criptografia
  - 4.4. Recursos para auditoria do sistema
    - 4.4.1. Desempenho do sistema
    - 4.4.2. Análise de erros
    - 4.4.3. Utilização de recursos físicos
    - 4.4.4. Acesso as bases de dados
    - 4.4.5. *Trace*
  - 4.5. Procedimentos de Instalação
    - 4.5.1. Rotinas para instalação do produto e configuração do ambiente
    - 4.5.2. Auto-explicação
    - 4.5.3. Acompanhamento contextual
  - 4.6. Conectividade
    - 4.6.1. Protocolo de rede
  - 4.7. Escalabilidade

**Figura 5.1 - Árvore específica de requisitos**

### **5.1.3.2. Elaboração do Questionário**

O questionário elaborado para os fornecedores dos produtos é apresentado no Apêndice II. Este instrumento contém questões objetivas e subjetivas, contemplando os requisitos presentes na árvore AER e alguns outros considerados importantes pelo *decision-maker*.



### 5.1.3.3. Classificação dos Grupos de Requisitos

A classificação dos requisitos presentes na AER em: essenciais, desejáveis e opcionais, representada nas Tabelas 5.1 (a), 5.1 (b) e 5.1 (c), respectivamente, levaram em consideração as prioridades da organização com base no levantamento realizado no item anterior.

<b>Requisitos Essenciais</b>
<b>1. Aspectos Gerais</b>
1.2.1. Guia do administrador de banco de dados
1.2.2. Guia de instalação e operação
1.2.5. Manual de referência do gerador de aplicação
1.2.6. Manual de referência do gerador de formulário
1.2.7. Guia de referência da linguagem de programação
1.2.9. Guia para o administrador da rede
1.2.11. Manual de referência da linguagem de consulta
1.3.1.2. Trimestral
1.4.1.3. Apenas em horário comercial, (07:00 às 22:00), em dias úteis
1.4.2.1. Telefone
<b>2. Ambiente de Desenvolvimento</b>
2.1.1. Integração com o Dicionário de Dados
2.2.2. Gerador de aplicação
2.2.3. Gerador de tela
2.2.4. Gerador de relatório
2.2.5. Gerador de menus
2.2.6. Gerador de gráficos
2.2.8. Manutenção do dicionário de dados
2.4. Independência entre dados e programa
2.5.1. Windows
<b>3. Administração dos Dados</b>
3.1.1. Inclusão de regras de validação
3.1.2. Consistência automática de campos
3.1.3. Referência cruzada a objeto
3.1.4. Geração automática de objeto
3.1.5. Mecanismos de distribuição do dicionário de dados
3.2.1. Hash
3.2.2. B-tree
3.3.1. Relacional
3.4.1. Tabelas
3.4.2. Colunas
3.4.3. Índices
3.4.4. Visões
3.4.5. Domínio
3.5.1. Queries
3.5.2. Joins
3.5.3. Projeções
3.5.4. Uniões
3.5.5. Interseções
3.5.6. Diferenças
3.6.1. Referencial
3.6.3. Domínio
3.6.4. Suporta triggers para outros tipos de integridade
3.7.1. Inteiro de dois bytes
3.7.2. Inteiro de quatro bytes
3.7.3. Inteiro de oito bytes
3.7.4. Real de ponto flutuante
3.7.5. Abstrato tipo data
3.7.6. Abstrato tipo money

3.7.7. Abstrato tipo boolean
3.7.8. Abstrato tipo time
3.8.1. Matemática
3.8.2. Estatística
3.8.3. Manipulação de strings
3.9. Otimização de consulta
<b>4. Administração do SGBD</b>
4.1. Controle de concorrência
4.2.2. Mecanismos de tolerância a falhas
4.2.3. Espelhamento de disco
4.2.1. Mecanismos para recuperação da base de dados
4.3.1. Recursos para definição de níveis de acesso
4.3.2. Criptografia
4.4.1. Desempenho do sistema
4.4.2. Análise de erros
4.4.3. Utilização de recursos físicos
4.4.4. Acesso as bases de dados
4.4.5. Trace
4.5.1. Rotinas para instalação do produto e configuração do ambiente
4.5.2. Auto replicação
4.5.3. Conservação dos dados do usuário
4.5.4. Acoplamento contextual
4.7. Escalabilidade

Tabela 5.1 (a) - Lista de requisitos essenciais

<b>Requisitos Desejáveis</b>
<b>1. Aspectos Gerais</b>
1.1. Existência de grupo de usuários
1.3.1.1. Mensal
1.4.2.2. Internet
1.4.2.3. Visita técnica
<b>2. Ambiente de Desenvolvimento</b>
2.1.2. Suporte a multimídia
2.1.3. Facilidade de uso (aspectos ergonômicos)
2.1.4. Disponibilidade de tutorial on-line
2.1.5. Suportar SQL embutido
2.2.1. Consulta ad-hoc
2.2.7. CASE
2.3.1. C
2.3.2. C++
2.3.5. Pascal
2.5.5. UNIX
<b>3. Administração dos Dados</b>
3.3.3. Orientado a objeto
3.6.2. Referencial declarativa

Tabela 5.1 (b) - Lista de requisitos desejáveis

<b>Requisitos Opcionais</b>
<b>1. Aspectos Gerais</b>
1.2.3. Guia do usuário para gerenciamento de objeto
1.2.4. Guia de release & conversão
1.2.8. Guia do usuário para o editor gráfico
1.2.10. Sumário de comandos da linguagem de consulta
1.3.1.3. Semestral
1.3.1.4. Anual
1.4.1.1. 24 horas apenas em dias úteis
1.4.1.2. 24 horas em úteis, feriados e finais de semana
<b>2. Ambiente de Desenvolvimento</b>
2.3.3. COBOL
2.3.4. Fortran
2.5.2. OS/2
2.5.3. VM
2.5.4. MVS

Tabela 5.1 (c) - Lista de requisitos opcionais



$$D_{ADD} = \frac{\sum_1^n d_i}{n} = \frac{0+1}{2} = 0,5$$

### Agregação dos Grupos da Classe de Requisitos Opcionais

$$O_{AG} = \frac{\sum_1^n o_i}{n} = \frac{1+1+1+1+1+1+0+0}{8} = 0,75$$

## Estágio 2 - Agregação de uma Classe de Requisitos

### Agregação dos Requisitos Essenciais

$$A_{RE} = \frac{\sum_1^m E_x}{m} = \frac{0+1+0+1}{4} = 0,5$$

### Agregação dos Requisitos Desejáveis

$$A_{RD} = \frac{\sum_1^m D_x}{m} = \frac{1+0,8+0,5+0}{3} = 0,43$$

### Agregação dos Requisitos Opcionais

$$A_{RO} = \frac{\sum_1^m O_x}{m} = \frac{0,75+0}{2} = 0,38$$

## Estágio 3 - Cálculo da “nota” a ser atribuída a cada SGBD

### Cálculo da “nota” do Sybase

$$N_{FINAL} = 6A_{RE} + 3A_{RD} + A_{RO} = 5,46 + 1,29 + 0,38 = 4,67$$

Um quadro resumo dos produtos classificados por notas obtidas através da análise dos requisitos quantificáveis é apresentado na Tabela 5.2.

Nome do Produto	Nota
Sybase	4,67
Oracle	3,95
Progress	3,15

**Tabela 5.2 - Lista de produtos classificados por notas**

### **Classificação Baseada nos Requisitos Não Quantificáveis**

Os resultados da desta análise são apresentados em dois estágios, descritos a seguir.

#### **Estágio 1 - Agregação dos requisitos não quantificáveis**

Neste estágio foi elaborada a lista dos requisitos não quantificáveis associados aos respectivos índices de satisfação atribuídos pelo *decision-maker*, a qual será apresentada a seguir na Tabela 5.3.

Requisito	Critério de Satisfação		
	Sybase	Oracle	Progress
<b>3. Dados sobre o Fornecedor e Fabricante</b>			
Início das Atividades do Fabricante	1	1	1
número de Produtos Lançados	1	1	1
número de fornecedores no País	2	1	1
Ano que iniciou a fornecer o produto	1	2	2
Número de Produtos Fornecidos	1	1	1
<b>2. Aspectos Gerais</b>			
Requisitos Mínimos de Hardware e Software	1	1	1
Ano de Lançamento do Produto no Brasil e no Exterior	1	1	1
Última Versão no Brasil e no Exterior	1	1	2
Quantidade de Cópias Vendidas no Brasil e no Exterior por Sistema Operacional	3	3	1
Lista de Aplicativos (comercializados) Compatíveis com o Produto	1	3	2
Existência de Grupos de Usuários	2	2	2
Relação de Cursos	1	3	2
Estados da Federação onde Normalmente são Oferecidos os Treinamentos	1	1	2
Duração Média dos Treinamentos	2	1	2
Número de Técnicos Disponível por Estado	2	1	2
Periodicidade de Atualização Tecnológica do Produto	2	1	2
<b>Administração do SGBD</b>			
Protocolos de Rede	1	1	1
Protocolos de Acesso a Bancos de Dados	1	1	1
Plataformas	1	1	1

**Tabela 5.3 - Índices de satisfação para os requisitos não quantificáveis**

#### **Estágio 2 - Classificação dos produtos baseada na análise dos requisitos não quantificáveis**

Neste estágio são apresentadas as pontuações alcançados por cada produto e a respectiva classificação, com base no atendimento aos requisitos não quantificáveis.

Total de Pontos alcançados pelo Sybase com base na expressão 4.8

$$TP1_{Pontos} = \sum_{i=1}^n r_i = 1+1+2+1+1+1+1+1+3+1+2+1+1+2+2+2+1+1+1 = 26$$

Total de Pontos alcançados pelo Oracle com base na expressão 4.8

$$TP2_{Pontos} = \sum_{i=1}^n r_i = 1+1+1+2+1+1+1+1+3+3+2+3+1+1+1+1+1+1 = 27$$

Total de Pontos alcançados pelo Progress com base na expressão 4.8

$$TP3_{Pontos} = \sum_{i=1}^n r_i = 1+1+1+2+1+1+1+2+1+2+2+2+2+2+2+2+1+1+1 = 29$$

Como pode ser observado na Tabela 5.4 o Sybase foi o produto que alcançou o melhor desempenho na análise dos requisitos não quantificáveis, ficando em segundo lugar o Oracle e em terceiro o Progress.

Classificação dos Produtos Baseada na Análise de Requisitos Quantificáveis	
Nome do Produto	Classificação
Sybase	Primeiro Lugar
Oracle	Segundo Lugar
Progress	Terceiro Lugar

Tabela 5.4 - Classificação dos produtos baseada na análise dos requisitos não quantificáveis

#### 5.1.4.2. Classificação Baseada nos Testes de Desempenho

Tendo em vista o cronograma de atividades para realização do estudo de caso com base no tempo disponível para a tomada de decisão pela empresa, o atraso na entrega dos questionários, por parte dos fornecedores, inviabilizou a realização desta análise.

#### 5.1.4.3. Classificação Baseada nos Custos

Só foi possível analisar os custos do *Sybase* e do *Oracle*, tendo em vista que o *Progress* não nos enviou a sua planilha de custos.

Requisito	Custos dos Produtos	
	Sybase	Oracle
SGBD	21.600,00	24.000,00
Módulos Adicionais	5.095,00	9.000,00
Suporte técnico (contrato anual)	3.600,00	2.500,00
Custo Total	30.295,00	35.500,00

**Tabela 5.5 - Planilha de custo dos produtos**

### **5.1.5 Classificação Baseada na Relação Custo Benefício**

Após uma avaliação criteriosa dos resultados obtidos através da aplicação da metodologia, o *decision-maker* selecionou o *Sybase* para ser adquirido por sua empresa. Tendo em vista o seu desempenho em todas fases do processo de seleção.

A partir dos resultados obtidos no Capítulo 2, que apontam para um maior índice de satisfação com os produtos instalados nas empresas que utilizaram algum método de seleção, e com base na análise das metodologias apresentadas no Capítulo 3, justificou-se o desenvolvimento de uma metodologia para apoiar o processo de seleção de SGBDs nas empresas.

### **6.1. Considerações Finais**

Segundo a empresa onde foi realizado o estudo de caso, apresentado no capítulo 5, apesar de não ter sido possível aplicar todas as fases da metodologia, esta contribuiu sobremaneira para o processo de seleção, ao atingir os seguintes objetivos:

- **Facilidade de aplicação:** as fases da metodologia foram definidas de forma concisa e de simples aplicação, baseando-se em cálculos matemáticos simples, que não demandam um treinamento específico para seu uso.
- **Avaliação criteriosa do produto:** com a aplicação da metodologia foi possível avaliar o produto no que diz respeito: (1) a compatibilidade dos recursos oferecidos com as necessidades da organização, (2) seu desempenho com base nas características de uma aplicação ou de conjunto de aplicações especificados pelo usuário, e (3) a relação custos associados e benefícios oferecidos.
- **Redução dos custos de seleção:** os custos associados ao processo de seleção são reduzidos, uma vez que pode ser aplicado pelos técnicos da própria empresa, que são quem melhor a conhecem.



## 6.2. Proposta para Trabalhos Futuros

Apesar de estar estruturada, ainda há aspectos a serem desenvolvidos para enriquecimento da metodologia apresentada neste trabalho, os quais citamos a seguir:

- Implementação de uma ferramenta para automatizar a metodologia.
- Elaboração de um guia de utilização descrevendo os passos da metodologia, e os conceitos utilizados.
- Expansão da árvore genérica de requisitos, de modo a incluir novos critérios de avaliação.
- Refinamento da fase análise custo benefício.
- Realização de um estudo sobre alternativas de testes de benchmark que possam ser incorporados à metodologia, por ocasião de sua automação.
- Expansão da metodologia para acomodar uma avaliação da influência do hardware sobre o desempenho dos produtos sob teste.

Espera-se que a Metodologia de Apoio à Decisão para Seleção de SGBD (MAPESGBD) se torne uma contribuição para as empresas usuárias de SGBD. Espera-se ainda que, de posse desta metodologia, o *decision-maker* passe a utilizar uma sistemática para apoiar o processo de seleção de um SGBD a partir das necessidades concretas das empresas, do perfil de seus usuários, de suas aplicações e do perfil geral da instalação.

## *Bibliografia*

- ANDERSON, T. L.; BERRE, A. J.; MALLILSON, M.; PORTER, H.; SHINEIDER, B. **The Tektronix HyperModel benchmark specification.** [S.l.,s.n.], 1989. (Tektronix Technical Report 8905).
- ANSI. American National Standards Institute. **The Database language SQL.** [S.l.,s.n.], 1986a. (Document. ANSI X9.135).
- ANSI. American National Standards Institute. **The Database language SQL.** [S.l.,s.n.], 1986b. (Document. ANSI X9.133).
- ANSI. American National Standards Institute. **The Database language SQL.** [S.l.,s.n.], 1989. (Document. ANSI X9.138).
- BUTLER, B. ; CANTER, B. SQL no PC: já se pode confiar neles?. **Pc Magazine Brasil**, v. 13, n. 12, p. 2-26, 1993.
- BELCHIOR, A. D. Aspectos tecnológicos envolvidos no desenvolvimento de software: um estudo de casos. In: SIMPÓSIO BRASILEIRO DE ENGENHARIA DE SOFTWARE, 9. , 1995, Recife. **Anais...** 1995. p. 311 – 326
- BITTON. D; DEWIT. D. J; TURBYFILL. C. **Benchmarking databases systems, a systematic approach.** In: PROC. OF THE VLDB COFERENCE, 1983, Florence. Italy.
- BITTON, D; HANRABAM, M. B; TURBYFILL. C. Performance of complex queries in main memory database systems. In: PROC. INTER. CONF. ON THIRD DATABASE ENGINEERING, 1987, Los Angeles.
- BAPTISTA, H. R; SIMÕES, L. J; GOYA, D. H; MADEIRA, M. A. SGBD programáveis. **PC Magazine Brasil**, v. 12, n.11, p. 39-74, 1992.
- BURNS, T. et al. Reference model for DBMS standardisation. **Sigmod Record**, v. 15, n. 1, p. 19-38, 1986.

- BOYD, H. W.; WESTFALL, K. **Pesquisa mercadológica: testes e custos**. 7.ed. Rio de Janeiro: Atlas, 1987.
- CASANOVA, M. A.; MOURA, A. V. **Princípios de sistemas de Gerência de bancos de dados distribuídos**. Rio de Janeiro: Campus, 1995. 355 p.
- CASTELL, R. G. G. **Object data management**. Califórnia: Addison Wesley, 1994. 389 p.
- CASTELL, R. G. G. **The object database standard: ODMG-93**. Califórnia: Morgan Kaufman, 1996. 184 p.
- CHEN, P. The entity relationship mode: toward united view of data. **TODS**, v. 1, p. 1, Mar. 1976.
- CODD, E. Extending the database relational model to capture more meaning. **TODS**, v. 4, p. 4, Dec. 1979.
- COBRA, M. **Administração de marketing**. São Paulo: Atlas, 1990.
- DATE, C. J. **Introdução a sistemas de bancos de dados**. Rio de Janeiro: Campus, 1991. 674 p.
- ELMASRI, R.; NAVATHE, S. B. **Fundamentals of database systems**. Redwood City: Adson Wesley, 1994. 873 p.
- FURTADO, A. L.; SANTOS, C. S. **Organização de banco de dados**. Rio de Janeiro: Campus, 1979.
- GANGER, G. R.; WORTHINGTON, B. L.; HOU, R. Y.; PATT, Y. N. *Disk arrays high performance, high reliability, storage subsystems*. In: **IEEE COMPUTER SOCIETY**, 1994. p. 30 - 36.
- GRADARIN, G.; VALDURIEZ, P. **Analysis and comparison of relational database systems**. Califórnia: Adison-Wesley, 1989.
- HANSON, R. J. **TPC Benchmark B - What it means and how to use it**. Disponível: AT&T Global Information Solutions (1996). <http://www.tpc.org/bdetail.html>. URL. (Consultado abril/1996).
- HAMMER, M.; MCLEOD, D. Database description with SDM: a semantic datamodel. **TODL**, v. 6, p. 3, Sept. 1981.
- KORTH, H. F.; SLBERSCHATZ, A. **Sistemas de banco de dados**. São Paulo: Makron, 1993. 748 p.

- LINTHICUM, D. S. The successes and failures of application development tools. **DBMS**, May. 1996. (<http://www.dbmsmag.com/9605d16.html>).
- LINTHICUM, D. S. Driving development. **DBMS**, Apr. 1997. (<http://www.dbmsmag.com/9704d14.html>).
- MARTINO, E. B. L. **Object-oriented database systems concepts and architectures**. Londres: Adilson-Wesley, 1994. 264 p.
- MARTIN, J. McCIURE, C. **Técnicas estruturadas e CASE**. São Paulo: McGraw-Hill, 1991. 854 p.
- MARTINS, L. M. F. **Um molde para o processo de produção de software comercial ilustrações com software UNIX de prateleira**, Campina Grande: UFPB-COPIN, 1993. 191 p. Dissertação Mestrado.
- PEREIRA, R. F, **Análise de desempenho de banco de dados utilizando benchmarks especializados**. Campinas: UNICAMP, 1990. Dissertação Mestrado.
- PEREIRA, R. M. P. **Análise de mercado de móveis do ponto de vista do consumidor: o caso da cidade de João Pessoa**. João Pessoa: UFPB, 1995. 150 p. Dissertação Mestrado.
- RAMALHO, J. A. ;FORESTI, N.; LAMOUNIER, A. Bancos de dados a aposta certa. **Informática EXAME Especial**, 1995.
- RUDIO, F. V. **Introdução ao projeto de pesquisa científica**. Petropolis: Vozes, 1978.
- SALEMI, J. **Guia para bancos de dados cliente/servidor**. Rio de Janeiro: Infobook, 1995. 343 p.
- SELLTIZ, L. M.; COOK, S. *Research methods in social relations*. New York: Kinehart and Winston, 1976.
- SHUMACHER, R. Talking toolsets. **DBMS**, Apr. 1996. (<http://www.dbmsmag.com/9704d15.html>).
- SMITH, J. SMITH, D. Database abstraction aggregation and generalisation. **TODS**, v. 2, p. 2. June 1977.
- SPIEGEL, M. R. **Estatística- resumo da teoria: 875 problemas resolvidos, 619 problemas propostos**. Brasilia: McGrawl-Hill, 1974. 580 p.

## *Apêndice I*

### ***Questionário Aplicado às Empresas Usuárias de SGBDs***

Neste apêndice, será apresentado o questionário aplicado na pesquisa descrita no Capítulo 2, com o objetivo de: (1) identificar os métodos e critérios utilizados para seleção de SGBDs, e (2) averiguar o nível de satisfação com relação aos recursos oferecidos pelos SGBDs (SGBDs).

**UFPB - Universidade Federal da Paraíba**  
**COPIN - Coordenação de Pós-graduação em Informática**

**Pesquisa:**

- 1. Dos métodos e critérios para seleção de SGBDs nas organizações**
- 2. Do nível de satisfação com relação aos recursos oferecidos pelos SGBDs (SGBDs)**

Adna Magnólia Dantas Silva  
Maria de Fatima Q. V. Turnell

João Pessoa, 28 de Dezembro de 1995

## Respondente

Pessoa Entrevistada :

.....

Cargo ou Função na Organização :

.....

Nome da Organização :

.....

.....

Endereço para Contato :

.....

.....

Perfil da Organização :

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Data da Entrevista : \_\_\_/\_\_\_/\_\_\_

Duração : \_\_\_\_\_

## Informações sobre os SGBDs Utilizados

SGBD 1 - Nome : .....  
Versão : .....  
Fabricante : .....  
Representante : .....  
Sistema Operacional: .....  
Hardware .....  
Data da Instalação : .....  
Modelo de dados : .....  
Utilização

- Frequente
- Ocasional
- Não Utilizado

SGBD 2 - Nome : .....  
Versão : .....  
Fabricante : .....  
Representante : .....  
Sistema Operacional: .....  
Hardware .....  
Data da Instalação : .....  
Modelo de dados : .....  
Utilização

- Frequente
- Ocasional
- Não Utilizado

SGBD 3 - Nome : .....  
Versão : .....  
Fabricante : .....  
Representante : .....  
Sistema Operacional: .....  
Hardware .....  
Data da Instalação : .....  
Modelo de dados : .....  
Utilização

- Frequente
- Ocasional
- Não Utilizado



SGBD 4 - Nome : .....  
Versão : .....  
Fabricante : .....  
Representante : .....  
Sistema Operacional: .....  
Hardware .....  
Data da Instalação : .....  
Modelo de dados : .....  
Utilização  
 Frequente  
 Ocasional  
 Não Utilizado

SGBD 5 - Nome : .....  
Versão : .....  
Fabricante : .....  
Representante : .....  
Sistema Operacional: .....  
Hardware .....  
Data da Instalação : .....  
Modelo de dados : .....  
Utilização  
 Frequente  
 Ocasional  
 Não Utilizado

### **Procedimentos para Escolha do SGBD**

1. Qual o SGBD mais utilizado?

.....  
.....  
.....  
.....  
.....  
.....

2. Quais os critérios utilizados para escolher o SGBD?

.....  
.....  
.....  
.....  
.....  
.....

3. Qual o método utilizado na escolha?

.....  
.....  
.....  
.....

4. Quais as fontes de informação (sobre o SGBD) utilizadas?

.....  
.....  
.....  
.....  
.....

5. Quem foi o responsável pela escolha?


.....

6. Qual o grupo de usuários do SGBD, sua empresa participa?

.....

### Matriz de Recursos X Nível de Satisfação / Tipo de Usuário de SGBD

Recursos Oferecidos pelo SGBD	Nível Externo		Nível Conceitual		Nível Interno	
	<i>Ad hoc</i>	Paramétrico	Projetista	Programador	Administrador	Analista
1. Linguagem de Consulta						
2. Gerador de Relatório						
3. Gerador de Formulário						
4. Dicionário de Dados						
5. Documentação						
6. Treinamento						
7. Suporte Técnico						
8. Ferramentas CASE						
9. Linguagem de Programação						
10. Aplicações Críticas						
11. Características dos Dados						
12. Recursos para Definição dos Níveis de Segurança						
13. Recurso para Alterar Esquema						
14. Métodos de Acesso aos Dados						
15. Mecanismos para Recuperação da Base de Dados						
16. Otimizador de Consulta						
17. Recursos para Auditoria do Sistema						
18. Adequação do SGBD com a arquitetura do ambiente						
19. Controle de Concorrência						
20. Integridade						
21. Portabilidade						
22. Interligação do Sistema Operacional com o SGBD						
23. Organização Física dos Dados						
24. Procedimentos de Instalação						
25. Nível Global de Satisfação						

 Não se aplica

---

## 1. Linguagem de Consulta

---

1.1. Linguagem utilizada :

.....  
.....

1.2. Nível de Satisfação :

	Excelente	Bom	Regular	Insatisfatório
1. Tempo de Resposta				
2. Interface (Facilidade de Uso)				
3. Qualidade da Apresentação da Informação				
4. Recursos de Hardware e Software				
5. Qualidade do Help on-line				

1.3. Dificuldades encontradas na geração de consultas :

.....  
.....  
.....  
.....

1.4. Qual o nível de integração com o dicionário de dados?

.....  
.....  
.....

1.5. Comentários

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

---

## 2. Gerador de Relatório

---

### 2.1. Gerador de relatório utilizado

.....  
.....

### 2.2. Facilidades encontradas na geração de relatórios

.....  
.....  
.....  
.....  
.....

### 2.3. Dificuldades encontradas na geração de relatórios

.....  
.....  
.....  
.....  
.....

### 2.4. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Interface (Facilidade de Uso)				
2. Qualidade da Apresentação da Informação				
3. Recursos de Hardware e Software				

### 2.5. Comentários

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

---

### 3. Gerador de Formulário

---

#### 3.1. Gerador de formulário utilizado

.....

#### 3.2. Facilidades encontradas na geração de formulários

.....  
.....  
.....  
.....

#### 3.3. Dificuldades encontradas na geração de formulários

.....  
.....  
.....  
.....

#### 3.4. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Interface (Facilidade de Uso)				
2. Qualidade da Apresentação da Informação				
3. Recursos de Hardware e Software				

#### 3.5. Comentários

.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....  
.....

---

## 4. Dicionário de Dados

---

### 4.1. Dificuldades encontradas na visualização dos dados

.....

.....

.....

.....

### 4.2. Como são os mecanismos de inclusão de regras de validação no dicionário de dados?

.....

.....

.....

### 4.3. A consistência de campos é feita automaticamente a partir de sua definição?

.....

.....

### 4.4. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Interface (Facilidade de Uso)				
2. Qualidade da Apresentação da Informação				

### 4.5. Comentários

.....

.....

.....

.....

.....

.....

.....

.....









---

## 8. Ferramentas CASE

---

8.1. Utiliza ferramenta CASE?

.....  
.....  
.....  
.....  
.....

9.2. Com que propósito?

- Análise
- Projeto
- Planejamento
- Geração de código

9.3. Comentários :

.....  
.....  
.....  
.....  
.....  
.....

---

## 9. Linguagem de Programação

---

9.1. Qual a linguagem de programação utilizada?

.....  
.....  
.....  
.....

9.2. Quais as dificuldades encontrada em programar na linguagem utilizada?

.....  
.....  
.....  
.....

9.3. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Facilidade de Uso				
2. Recursos disponíveis				

9.4. Comentários :

.....  
.....  
.....  
.....  
.....  
.....

---

## 10. Aplicações Críticas

---

10.1 . Descreva sucintamente a sua aplicação

.....  
.....  
.....  
.....

10.2. Dificuldades encontradas na interação com a aplicação

.....  
.....  
.....  
.....  
.....

10.3..O programa de aplicação atende as suas necessidades

.....  
.....

10.4. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Tempo de Resposta				
2. Facilidade de Uso				
3. Qualidade da Apresentação da Informação				
4. Recursos de Hardware e Software				

10.5. Comentários

.....  
.....  
.....  
.....



---

## 12. Recursos para Definição de Níveis de Segurança

---

12.1. Quais os níveis de acesso aos dados? (por objetos: tabelas, colunas, usuários e por funções: insert, update, delete, select).

.....  
.....  
.....

12.2. Os dados podem ser criptografados? Quais as facilidades e dificuldades do procedimento?

.....  
.....  
.....  
.....

### 12.3. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Facilidade de Uso				
2. Qualidade da Apresentação da Informação				
3. Mecanismos para Auditoria				

### 12.4. Comentários

.....  
.....  
.....  
.....

---

### 13. Recursos para Alterar Esquema

---

13.1. A alteração da definição de um esquema implica na recompilação dos programas que o utilizam?

.....

13.2. Recursos disponíveis para alterar esquema

.....

.....

.....

13.3. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Facilidade de Uso				

13.4. Comentários

.....

.....

.....

.....

.....



---

## 14. Métodos de Acesso aos Dados

---

14.1. Quais os métodos de acesso permitidos ? (por índices, chave primária, chave secundária, hash, B-tree).

.....  
.....  
.....

14.3. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Facilidade de Uso dos Recursos				

14.3. Comentários

.....  
.....  
.....

---

## 15. Mecanismos para Recuperação da Base de Dados

---

15.1. Existem mecanismos de back-up e restore?

.....

15.2. É possível a realização de back-up on-line?

.....

.....

15.3. Existem mecanismos de tolerância e recuperação de falha?

.....

.....

15.4. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Facilidade de Uso dos Recursos				

15.3. Comentários

.....

.....

.....

.....

---

## 16. Otimização de Consulta

---

16.1. Qual o tipo de otimização utilizada

- Baseada em custo
- Baseada em sistema
- Dinâmica
- Distribuída

16.2. Interação com o otimizador

.....

.....

.....

.....

16.3. Nível de Satisfação

	Excelente	Bom	Regular	Insatisfatório
1. Facilidade de Uso				

16.4. Comentários

.....

.....

.....

.....

.....



---

## 18. Adequação do SGBD com a Arquitetura do Sistema

---

### 18.1. Arquitetura

- Centralizada
- Cliente/Servidor
- Distribuída

### 18.2. Adequação do SGBD com relação a arquitetura dos recursos da organização

.....

.....

.....

.....

.....

.....

.....

.....

## Apêndice II

### ***Questionário Aplicado às Empresas Fornecedoras de SGBDs***

Neste apêndice é apresentado o questionário elaborado pela empresa piloto, quando da aplicação da MAPESGBD, com o objetivo de avaliar os fornecedores de SGBDs participantes do processo de seleção e seus produtos.

**Levantamento de Características  
do SGBD**

**Empresa Fornecedora :** \_\_\_\_\_

**Endereço do Fornecedor :** \_\_\_\_\_

**Telefone/Fax :** \_\_\_\_\_

**Nome do Respondente:** \_\_\_\_\_

**Cargo do Respondente :** \_\_\_\_\_

\_\_\_\_\_  
**Assinatura do Respondente**

**Obs.: As páginas seguintes deverão ser rubricadas.**

## Questionário ao Fornecedor

### 1. Dados sobre o Fornecedor e Fabricante

#### 1.1. Identificação do Fabricante

Nome \_\_\_\_\_

Endereço : \_\_\_\_\_

Data de Fundação : \_\_\_\_\_

#### 1.2. Dados sobre o Fabricante

Início de atividade : \_\_\_\_\_

Número de produtos lançados : \_\_\_\_\_

Número de fornecedores no país : \_\_\_\_\_

#### 1.3. Dados sobre o Fornecedor

Ano que iniciou a fornecer o produto : \_\_\_\_\_

Número de produtos fornecido : \_\_\_\_\_

### 2. Aspectos Gerais sobre o Produto

#### 2.1. Requisitos mínimos de hardware e software

Recurso mínimo de memória auxiliar e principal para instalação da configuração completa

\_\_\_\_\_

Recurso mínimo de memória auxiliar e principal para instalação da configuração mínima

\_\_\_\_\_

Memória principal requerida pelo próprio código objeto do sistema para as áreas de comunicação (buffers)

\_\_\_\_\_

Memória necessária para armazenar estruturas auxiliares, elos, índices, inversões, etc.

\_\_\_\_\_

CPU recomendada

\_\_\_\_\_

2.2. Ano de lançamento do produto no Brasil e no exterior : \_\_\_\_\_

2.3. Última versão no Brasil e no exterior : \_\_\_\_\_



2.4. Quantidade de cópias vendidas por sistema operacional no Brasil e no exterior :

---

---

---

---

2.5. Lista de aplicativos (comercializados) compatíveis com o produto

---

---

---

---

2.6. Existência de grupos de usuários

- Sim
- Não

2.7. Documentação disponível

- Guia do administrador de banco de dados
- Guia de instalação e operação
- Guia do usuário para gerenciamento de objeto
- Guia de *release* & conversões
- Manual de referência do gerador de aplicação
- Manual de referência do gerador de formulários
- Guia de referência de linguagem de programação
- Guia do usuário para o editor gráfico
- Guia para o administrador da rede
- Sumário de comandos da linguagem de consulta
- Manual de referência da linguagem de consulta
- Manual de códigos e mensagens de erros
- Outros \_\_\_\_\_

2.8. Treinamento

Periodicidade dos treinamentos

- Mensal
- Trimestral
- Semestral
- Anual
- Outros \_\_\_\_\_

Relação de cursos

---

---

Estados da Federação onde normalmente são oferecidos os treinamentos

---

---

Duração média dos treinamentos

---

---

## 2.9. Suporte técnico

### Atendimento

- 24 horas apenas em dias úteis
- 24 horas em dias úteis, feriados e finais de semana
- Apenas em horário comercial (07:00 às 22:00), em dias úteis
- Outros \_\_\_\_\_

### Formas de atendimento

- Telefone
- Internet
- Visita técnica
- Outros : \_\_\_\_\_

Número de técnicos disponíveis por Estado : \_\_\_\_\_

2.10. Periodicidade de atualização tecnológica do produto : \_\_\_\_\_

## 3. Ambiente de Desenvolvimento

### 3.1. Linguagem de consulta

- Integração com o dicionário de dados
- Suporte a multimídia
- Facilidade de uso (aspectos ergonômicos)
- Disponibilidade de tutorial *on-line*
- Implementação SQL padrão ANSI

### 3.2. Ferramentas disponíveis

- Consulta ad-hoc
- Gerador de aplicação
- Gerador de telas
- Gerador de relatórios
- Gerador de menus
- Gerador de gráficos
- CASE
- Manutenção do Dicionário de Dados
- Outros : \_\_\_\_\_

### 3.3. Linguagem de programação

- COBOL
- C
- C ++
- Pascal
- Fortran
- Outros \_\_\_\_\_

### 3.4. O produto garante a independência entre o código gerado pelos programas e os dados manipulados?

- Sim
- Não

### 3.5. Portabilidade de código

- Windows
- OS/2
- VM
- MVS
- UNIX
- Outros \_\_\_\_\_

## 4. Administração dos Dados

### 4.1. Dicionário de dados

- Inclusão de regras de validação
- Consistência automática de campos
- Referência cruzada a objeto
- Geração automática de objetos
- Mecanismos de distribuição do dicionário de dados

### 4.2. Métodos de acesso aos dados

- Índice
- Hash
- B-tree
- Outros \_\_\_\_\_

### 4.3. Modelo de dados

- Relacional
- Relacional estendido
- Orientado a objeto
- Outros \_\_\_\_\_

### 4.4. Estruturas de informação relacionais

- Tabelas
- Colunas
- Índices
- Visões
- Domínios

4.5. Operações sobre objetos implementadas pelo produto

- Queries*
- Joins*
- Projeções
- Uniões
- Interseções
- Diferenças
- Outros \_\_\_\_\_

4.6. Quanto as limitações em relação a:

Número máximo de tabelas por base de dados: \_\_\_\_\_

Ou número máximo de linhas por tabela: \_\_\_\_\_

Ou número máximo de colunas por tabela: \_\_\_\_\_

Ou Número máximo de índices por tabela: \_\_\_\_\_

Ou número máximo de colunas por índice: \_\_\_\_\_

Ou número máximo de colunas por chave composta: \_\_\_\_\_

Ou tamanho máximo de colunas alfanuméricas: \_\_\_\_\_

Ou valor máximo de colunas numéricas: \_\_\_\_\_

Ou outros: \_\_\_\_\_

4.7. Tipos de integridade de dados implementada:

- Referencial
- Referencial declarativa
- Domínio
- Suporta *triggers* para outros tipos de integridade
- Outros \_\_\_\_\_

4.8. Tipos de dados suportados pela linguagem de definição de dados:

- Inteiro de dois bytes
- Inteiro de quatro bytes
- Inteiro de oito bytes
- Real de ponto flutuante
- Abstrato tipo data
- Abstrato Tipo *Money*
- Boolean*
- Time*
- Multimídia
- Outros \_\_\_\_\_

4.9. A linguagem de manipulação de dados oferece funções:

- Matemáticas
- Estatísticas
- Manipulação de strings
- Outros \_\_\_\_\_

4.10. O produto implementa otimizador de consulta

- Sim
- Não

## 5. Administração do SGBD

5.1. O produto oferece controle de concorrência no acesso aos dados

- Sim
- Não

5.2. Recursos para assegurar a integridade física dos dados

- Mecanismos para recuperação da base de dados
- Mecanismos de tolerância a falha
- Espalhamento de arquivos
- Outros \_\_\_\_\_

5.3. Recursos para controle de acesso

- Recursos para definição de níveis de acesso
- Processo para Criptografia
- Outros \_\_\_\_\_

5.4. Recursos para auditoria do sistema

- Desempenho do sistema
- Análise de erros
- Utilização de recursos físicos
- Acesso as bases de dados
- Trace*
- Outros \_\_\_\_\_

5.5. Procedimentos de Instalação oferecem:

- Rotinas para instalação do produto e configuração do ambiente
- Auto explicação
- Conservação dos dados do usuário
- Acompanhamento contextual
- Outros \_\_\_\_\_

5.6. Conectividade

Protocolos de rede

\_\_\_\_\_  
\_\_\_\_\_

Protocolo de acesso a banco de dados

\_\_\_\_\_  
\_\_\_\_\_

5.7. Escalabilidade

Não

Sim

Plataformas

---

---

**6. Estimativa de Custo**

6.1. Preço da configuração mínima

Descrição do Módulo	R\$

6.2. Preço de cada de módulo adicional

Descrição do Módulo	R\$

6.3. Custos adicionais (treinamento, instalação, suporte, documentação adicional, outros)

Descrição	R\$

