

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CURSO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

**ESTUDO E IMPLEMENTAÇÃO DE UM SISTEMA DE
BANCO DE DADOS DISTRIBUÍDO**

Dissertação submetida à Universidade Federal de Santa Catarina para
a obtenção do grau de Mestre em Ciência da Computação.

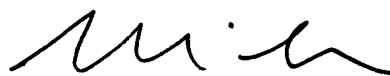
Nery Artur Eller

Florianópolis, Abril de 1997

ESTUDO E IMPLEMENTAÇÃO DE UM SISTEMA DE BANCO DE DADOS DISTRIBUÍDO

Nery Artur Eller

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, especialidade Sistemas de Computação e aprovada em sua forma final pelo programa de Pós-Graduação em Ciência da Computação.



Prof. Dr. Murilo Silva de Camargo
Orientador e Coordenador do Curso

Banca Examinadora :



Prof. Dr. Murilo Silva de Camargo
Presidente



Prof. Dr. Vitorio B. Mazzola



Prof. M.Sc. Fábio Teixeira de Campos



Prof. M.Sc. Elizabeth Sueli Specialski

Agradecimentos

Ao orientador Murilo Silva de Camargo pela dedicação, competência e tenacidade demonstradas na orientação deste trabalho, além do incentivo e amizade ao longo de nossa breve convivência.

Ao professor Fábio Teixeira de Campos, pelo grande apoio prestado ao longo deste trabalho.

A professora Elizabeth Sueli Specialski, pelo tratamento prestativo, educado e eficiente.

Ao prof. Vitório B. Mazzola, pela sua importante participação na banca examinadora.

A minha família pelo apoio.

Em especial à minha esposa Eliane Schürhaus Eller, pela compreensão, paciência e incentivos.

A Universidade Federal de Santa Catarina, que tornou possível esta importante conquista.

SUMÁRIO

LISTA DE ILUSTRAÇÕES	vi
LISTA DE SIGLAS	vii
RESUMO	viii
ABSTRACT	ix
1 - INTRODUÇÃO	1
2 - BANCO DE DADOS DISTRIBUÍDO	4
2.1 - Introdução	4
2.2 - Conceitos básicos	5
2.3 - Descrição de banco de dados distribuído	6
2.3.1 - Requisitos funcionais de um banco de dados distribuído ideal	6
2.3.2 - Arquitetura de referência para um banco de dados distribuído	8
2.3.3 - Projeto de banco de dados distribuído	9
2.4 - Sistemas de gerência de BDD's	14
2.4.1 - Arquitetura de um SGBDD	14
2.4.2 - Gerência de transações distribuídas	15
2.4.3 - Gerência de controle de concorrência	17
2.4.4 - Gerência de recuperação	21
2.4.5 - Integridade e segurança	25
2.4.6 - Administração de bancos de dados distribuídos	26
2.5 - Conclusão	29
3 - SISTEMA DE CONTROLE DA FROTA DE VEÍCULOS	31
3.1 - Introdução	31
3.2 - Apresentação do sistema	31
3.3 - Descrição do sistema	32
3.3.1 - Diagrama de contexto	33
3.3.2 - Diagrama de fluxo de dados de nível 0	34
3.3.3 - Diagrama E-R	37
3.3.4 - Diagrama do fluxo de informações	39
3.3.5 - Grade das ocorrências anuais das principais tabelas	40
3.3.6 - Descrição dos módulos do sistema	41
3.4 - Características particulares (processos distribuídos)	46
3.5 - Principais aplicativos	47
3.6 - Conclusão	50
4 - DESCRIÇÃO DO AMBIENTE DE IMPLEMENTAÇÃO	52
4.1 - Introdução	52
4.2 - Ambiente de implementação	53

4.3 - Principais características do SGBD Oracle7	54
4.4 - Mecanismos para distribuição de dados no Oracle7	59
4.5 - Conclusão	64
5 - PROJETO LÓGICO DISTRIBUÍDO	65
5.1 - Introdução	65
5.2 - Matriz processos x locais	67
5.3 - Matriz processos x assuntos de informação	68
5.4 - Matriz assuntos de informação x locais	69
5.5 - Plano de compartilhamento lógico	70
5.6 - Conclusão	72
6 - PROJETO FÍSICO DISTRIBUÍDO	73
6.1 - Introdução	73
6.2 - Assuntos de distribuição	74
6.3 - Definição dos locais de distribuição	76
6.4 - Definição da forma de atualização	77
6.5 - Distribuição Física dos assuntos de distribuição	78
6.6 - Especificação de distribuição	79
6.7 - Plano de compartilhamento físico	80
6.8 - Detalhes de implementação	81
6.9 - Conclusão	84
7 - ANÁLISES DO MODELO IMPLEMENTADO	85
7.1 - Introdução	85
7.2 - Análise do modelo implementado com relação ao modo centralizado	86
7.2.1 - Tempo de resposta	86
7.2.2 - Transmissão de dados na rede	86
7.2.3 - Armazenamento	88
7.2.4 - Disponibilidade	89
7.2.5 - Manutenção de sistemas	89
7.2.6 - Desenvolvimento e documentação de sistemas	90
7.2.7 - Backup e recuperação	90
7.2.8 - Funções do DBA	91
7.2.9 - Informações gerenciais	91
7.3 - Análise do modelo implementado quanto ao banco de dados ideal	92
7.3.1 - Autonomia local	92
7.3.2 - Sem dependência de um centro de proc. de dados central	92
7.3.3 - Operação contínua	93
7.3.4 - Independência local	93
7.3.5 - Independência de fragmentação	93
7.3.6 - Independência de replicação	94
7.3.7 - Processamento de consulta distribuída	94
7.3.8 - Gerenciamento de transações distribuídas	94

7.3.9 - Independência de máquina	95
7.3.10 - Independência do sistema operacional	95
7.3.11 - Independência de redes de comunicação	95
7.3.12 - Independência de SGBD	96
7.4 - Conclusão	96
8 - CONCLUSÃO	97
ANEXOS	100
REFERÊNCIAS BIBLIOGRÁFICAS	106

LISTA DE ILUSTRAÇÕES

- 1 - Representação da arquitetura de referência para um banco de dados Distribuído
- 2 - Fragmentação vertical seguida de uma fragmentação horizontal
- 3 - Fragmentação horizontal seguida de uma fragmentação vertical
- 4 - Réplica periódica
- 5 - Réplica contínua
- 6 - Modelo referencial de recuperação de transação distribuída
- 7 - Grafo wait-for de detecção de bloqueio mútuo
- 8 - Árvore de detecção de bloqueio mútuo
- 9 - Detector de bloqueio global
- 10 - Diagrama de Contexto do sistema de informação
- 11 - Diagrama do Fluxo de Dados de nível 0 (parte 1) do sistema de informação
- 12 - Diagrama do Fluxo de Dados de nível 0 (parte 2) do sistema de informação
- 13 - Diagrama E-R do sistema de informação
- 14 - Diagrama do Fluxo de Informações do sistema implementado
- 15 - Planilha das ocorrências das tabelas do sistema de informação
- 16 - Topologia da rede de comunicação do ambiente de implementação
- 17 - Protocolo Bifásico de Confirmação
- 18 - Exemplo de uso da Categoria de Associação
- 19 - Assunto de Distribuição Acidentes

LISTA DE SIGLAS

BUV : Boletim de Utilização de Veículo

DA : Administrador de Dados.

DBA : Administrador de Banco de Dados.

DDL : Linguagem de Definição de Dados.

DML : Linguagem de Manipulação de Dados.

LAN : Rede Local de Comunicação de Dados.

ODBC : Open Database Connectivity.

DRDA : Distributed Relational Database Architecture.

OSV : Ordem de Serviço com Veículo.

SGBD : Sistema Gerenciador de Banco de Dados.

SGBDD : Sistema Gerenciador de Banco de Dados Distribuído.

SQL : Linguagem Estruturada de Consulta.

WAN : Rede de Longa Distância de Comunicação de Dados.

2PC : Protocolo Bifásico de Confirmação.

2PL : Protocolo Bifásico de Bloqueio.

RESUMO

Neste trabalho apresenta-se o desenvolvimento e implementação de um sistema de informação baseado em um Banco de Dados Distribuído. No desenvolvimento do sistema utilizou-se técnicas da Engenharia da Informação e métodos de distribuição de dados para realizar o levantamento dos dados e processos e também os projetos lógico e físico. A implementação do sistema foi baseada no Sistema de Gerência de Banco de Dados Distribuído Oracle e em uma grande variedade de outros recursos computacionais da Celesc - Centrais Elétricas de Santa Catarina S.A, empresa usuária do sistema desenvolvido. No desenvolvimento do sistema deu-se ênfase às técnicas de replicação dos dados e do mecanismo de manutenção da integridade do Banco de Dados Distribuído em operações de atualizações globais, visando a garantia da consistência dos dados. Tirando proveito da experiência adquirida na implementação desse sistema, realizou-se também neste trabalho uma análise comparativa dos principais requisitos ideais de Bancos de Dados Distribuídos existentes na literatura com o sistema implementado. Outrossim, um importante aspecto prático deste trabalho foi o domínio de modernas tecnologias para desenvolvimento de sistemas de informação distribuídos para dotar e repassar a Celesc uma nova metodologia de desenvolvimento de sistemas, atendendo com isto necessidades emergentes da empresa nesta área.

ABSTRACT

The goal of this work is to present the development and implementation of an information system based on a distributed database. Techniques of Information Engineering were used in the development of the system as well as data distribution methods for the data surveying and processing and for the logical and physical projects. The system implementation was based on the Oracle Distributed Database Management and a great variety of Celesc's - Centrais Elétricas de Santa Catarina S.A - computing resources, company user of the developed system. A special emphasis was given to data replication techniques in the development of the system, as well as the mechanism of integrity maintenance of the distributed database in the global updating operations, focusing on the guarantee of the data consistency. Capitalizing from the implementation of this system, a comparative analysis of the main ideal requirements of distributed database existing literature with the implemented system was also made. An important practical aspect of this work was the mastering of modern technologies for the development of distributed data systems to equip Celesc with a new system development methodology answering the most urgent needs of the company in this area.

CAPÍTULO 1

INTRODUÇÃO

Nos últimos anos os Sistemas de Gerência de Bancos de Dados têm desempenhado um papel fundamental para o desenvolvimento de sistemas de informação. Os sistemas são, em sua maioria, implementados de forma centralizada, motivadas por razões econômicas, técnicas e culturais. Por outro lado, com o avanço das técnicas de comunicação de dados, aliado ao desenvolvimento e barateamento dos hardwares, foi possível migrar os sistemas de informação da tradicional forma centralizada para uma arquitetura distribuída, fazendo-se uso da tecnologia de Banco de Dados Distribuído.

A Celesc, empresa distribuidora de energia elétrica para todo o estado catarinense, sente necessidade do domínio desta tecnologia para que possa fazer uso na migração de seus sistemas de informação para a arquitetura distribuída.

Neste trabalho foi desenvolvido e implementado um sistema de informação distribuído que tem como uma de suas finalidades dotar e repassar a Celesc um conhecimento abrangente nesta área.

Esta implementação foi confrontada com os requisitos básicos de um Banco de Dados Distribuído e também quanto à forma centralizada do sistema de informação implementado.

Ao se confrontar o sistema de informação distribuído implementado com o modelo centralizado, pretende-se levantar os benefícios auferidos com o seu uso e, desta forma, justificar a sua adoção na empresa. O confronto entre as características do banco de dados distribuído implementado e as consideradas ideais darão uma idéia de quanto os atuais Sistemas Gerenciadores de Banco de Dados Distribuídos terão que ser aperfeiçoados. Poderá servir também como

subsídio para se identificar a lacuna entre o projeto lógico do sistema distribuído traçado e o projeto físico implementado.

Por que a Celesc, já possuindo um SGBDD que implementa algumas características de bancos de dados distribuídos, não faz uso desta tecnologia para integrar os diversos sistemas de informação existentes? O presente trabalho é proposto justamente para responder parcialmente esta questão.

Neste trabalho realizou-se um estudo sobre bancos de dados distribuídos e sobre o SGBDD Oracle7 (versão atual) e, a partir deste estudo, desenvolveu-se e implementou-se um sistema de informação distribuído. O sistema escolhido foi o Sistema de Gerência da Frota de Veículos da Celesc. Na escolha do sistema, levou-se em consideração o seu porte, para tornar possível a implementação; as suas características distribuídas a nível de processos e dados; o nível da área usuária, para se efetuar um bom levantamento de dados e processos e, por último, sua abrangência com relação ao número de sítios envolvidos.

Mesmo sabendo antecipadamente dos problemas relacionados às limitações da estrutura de comunicação de dados; da inexistência de metodologia interna que contemple esse assunto; das novas atribuições dos DBA's e DA e de algumas limitações apresentadas pelo SGBDD Oracle7 com relação à distribuição de dados, achou-se importante a realização deste trabalho para prover a empresa de uma nova tecnologia de desenvolvimento de sistemas. Atualmente, os processos gerenciais da Celesc são bastante centralizados, porém, aos poucos, a descentralização administrativa está ocorrendo. E, acompanhando este processo, está a idéia do uso de banco de dados distribuído na empresa.

Outrossim, do ponto de vista acadêmico, o desenvolvimento de um Sistema de Banco de Dados Distribuído é uma experiência relevante, no sentido de se poder confrontar as limitações práticas impostas pelas restrições de implementação com as características técnicas exigidas em sistemas desse tipo.

O trabalho está organizado da forma que segue:

O capítulo 2 é dedicado aos estudos efetuados sobre Bancos de Dados Distribuídos. De uma forma resumida é descrita a gerência de transações distribuídas, gerência de concorrência, gerência de recuperação e o protocolo bifásico de confirmação (2PC); integridade e segurança de dados em um ambiente distribuído; catálogo e dicionário de dados e, finalmente, funções e atribuições do DBA na administração de um ambiente distribuído.

O sistema de informação distribuído que serviu como objeto de estudo é detalhado no capítulo 3. O sistema, Controle da Frota de Veículos, é descrito através de seu Diagrama de Contexto, Diagrama de Fluxo de Dados de Nível 0, Diagrama de Entidade-Relacionamento e Diagrama do Fluxo de Informações.

Descreve-se também em detalhes os módulos do sistema, os principais processos distribuídos e aplicativos.

O ambiente de implementação do sistema de informação distribuído é apresentado no capítulo 4. A descrição do SGBDD Oracle7 toma a maior parte deste capítulo em virtude da sua importância no desenvolvimento deste trabalho e também das análises apresentadas no capítulo 7.

O capítulo 5 aborda o Projeto Lógico Distribuído. No desenvolvimento do projeto lógico constrói-se uma série de matrizes que culminará com o Plano de Compartilhamento Lógico, relacionando em cada local físico de abrangência do sistema a forma de distribuição dos assuntos de informação (subsistemas) [9].

O Projeto Físico Distribuído é descrito no capítulo 6. Para o desenvolvimento do projeto físico distribuído utiliza-se como insumo o Plano de Compartilhamento Lógico gerado no Projeto Lógico. O Projeto Físico indicará a forma de distribuição de cada tabela do sistema em cada local físico de abrangência do sistema [9].

No capítulo 7 são realizadas 2 análises: a primeira analisa o modelo de banco de dados distribuído implementado com relação ao mesmo modelo desenvolvido de forma centralizada; a segunda análise vincula o modelo de banco de dados distribuído implementado e os aspectos funcionais ideais deste.

Finalizando este trabalho, são tecidas algumas considerações conclusivas obtidas ao longo desse estudo, apresentando-se algumas sugestões para o aperfeiçoamento do trabalho em função de melhorias a serem realizadas no atual ambiente de implementação.

Capítulo 2

BANCO DE DADOS DISTRIBUÍDO

2.1 - Introdução

Neste capítulo apresenta-se um estudo sobre Banco de Dados Distribuído. Abordam-se detalhes de Banco de Dados Distribuído e também as principais características e funções desses.

Neste resumo da teoria de Banco de Dados Distribuído foram selecionados alguns parâmetros que servirão de base às análises do sistema de informação distribuído implementado que é objeto deste trabalho.

Sobre Bancos de Dados Distribuídos, destacam-se o projeto de banco de dados distribuído e os seus componentes : projeto de fragmentação, projeto de replicação e projeto de alocação de dados. O principal enfoque na implementação do sistema Controle da Frota de Veículos é o projeto de replicação, totalmente suportado pelo SGBDD Oracle7 [17]. O projeto de fragmentação não é suportado pelo SGBDD Oracle7. O projeto de alocação neste trabalho é constituído basicamente do levantamento das tabelas, réplicas e índices destinados a cada sítio e pelos cálculos efetuados para se chegar ao espaço em disco necessário para armazená-los.

Os Sistemas de Gerência de Banco de Dados Distribuídos são estudados através de seus principais componentes : gerência de transação, concorrência e recuperação. O método mais utilizado pelos SGBD's comerciais para o controle de concorrência é o Bloqueio. Para manter a integridade dos dados envolvidos em transações distribuídas, a maioria dos SGBD's utiliza o Protocolo Bifásico de

Confirmação (2PC). Para resolver os casos de deadlocks globais, o método geralmente implementado é o Timeout [4].

Para a construção deste capítulo utilizou-se como fonte básica de pesquisa os livros [4] e [6].

2.2 - Conceitos Básicos

Antes de se iniciar o estudo sobre bancos de dados distribuídos serão introduzidos os conceitos de banco de dados, sistema gerenciador de banco de dados, banco de dados distribuído e sistema gerenciador de banco de dados distribuído, já que esses assuntos serão bastante discutidos durante o desenvolvimento deste trabalho.

Banco de Dados (BD) : Um banco de dados pode ser definido como uma coleção computadorizada de dados operacionais armazenados que servem para suprir as necessidades de múltiplos usuários dentro de uma ou mais organizações [4].

Sistema Gerenciador de Banco de Dados (SGBD) : De uma forma bastante simples pode-se dizer que o SGBD é um conjunto de softwares responsável pelo controle e gerenciamento dos recursos de banco de dados [12].

Banco de Dados Distribuído (BDD) : Um banco de dados distribuído é uma coleção de dados pertencentes logicamente ao mesmo sistema, mas distribuídos sobre vários sítios de uma rede de computadores [16].

Para Christian J. Date “Um sistema de banco de dados distribuído consiste de uma coleção de sítios, conectados através de redes de comunicação, no qual cada sítio é um sistema de banco de dados em seu próprio direito, mas os sítios concordam em cooperar, de tal forma que um usuário em qualquer sítio possa acessar qualquer dado na rede exatamente como se estivesse armazenado no próprio sítio do usuário.” [10]

2.3 - Descrição de Banco de Dados Distribuído

2.3.1 - Requisitos Funcionais de um Banco de Dados Distribuído Ideal

Segundo Christian J. Date um banco de dados distribuído ideal deve atender os seguintes requisitos:

Autonomia Local : Os sítios, em um sistema distribuído, devem ser autônomos, ou seja, os dados são acessados e gerenciados localmente, as operações locais permanecem no mesmo sítio e nenhum sítio depende de outro para se tornar operável. Não deve existir dicionário centralizado, controlador de atividades centralizado ou detector de bloqueio mútuo centralizado.

Independência dos Sítios : Todos os sítios em um sistema distribuído devem possuir o mesmo tratamento. As funções de gerenciamento de dicionário, processamento de pesquisas, controle de concorrência e recuperação não devem depender de um sítio central.

Operação Contínua : A entrada de um novo sítio, instalação de uma nova versão do SGBD são exemplos de operações que não devem alterar em nada o funcionamento do sistema ou das aplicações. Cada sítio é uma unidade funcional independente.

Independência de Local : Os dados aparecem para os usuários como se estivessem armazenados localmente. A movimentação de dados entre sítios não deve alterar a forma do usuário ver os dados. O usuário deve consultar e atualizar dados independente do sítio.

Independência de Fragmentação : Fragmentação é a divisão de uma relação utilizando-se operações relacionais de restrições e projeções. Esta divisão pode ocorrer de forma horizontal, envolvendo subconjuntos de linhas, de forma vertical, envolvendo subconjuntos de colunas e de forma mista. A reconstrução da relação original dos fragmentos é feita utilizando-se operadores de junção e de união da Linguagem de Manipulação de Dados (SQL). A visão que um usuário possui de uma relação fragmentada é a mesma que ele possui de uma relação centralizada.

Independência de Replicação : Um sistema de banco de dados distribuído deve suportar replicação de dados, ou seja, deve permitir que uma relação (ou fragmento) possua uma ou mais cópias em vários locais distintos da rede. As razões básicas de uso de réplicas são a disponibilidade dos dados e o desempenho das aplicações. A localização e manipulação das réplicas de dados devem ser transparentes ao usuário.

Processamento de Consultas Distribuídas : A otimização de consultas em um sistema distribuído deve ser também distribuída, envolvendo normalmente uma etapa de otimização global para minimizar o tráfego na rede, seguido das etapas seguintes de otimização local em cada sítio envolvido.

Gerenciamento de Transações Distribuídas : Dois aspectos importantes no gerenciamento de uma transação distribuída são o controle de concorrência e o controle de recuperação. Em uma transação distribuída, a atomicidade da transação deve ser obtida, da mesma forma como ocorre em um sistema centralizado. O método geralmente utilizado para manter a integridade dos dados em uma transação distribuída é conhecido como Protocolo Bifásico de Confirmação. Os SGBDs comerciais, em sua maioria, utilizam o bloqueio de linha para atualizações e, para detectar bloqueios mútuos globais, o mecanismo de expiração de tempo.

Independência de Máquina : Os SGBDD's utilizados em um sistema de banco de dados distribuídos devem suportar qualquer plataforma de hardware, ou seja, o fato de rodar em uma ou outra máquina deve ser transparente ao sistema.

Independência do Sistema Operacional : O sistema operacional que executará o SGBDD deve ser transparente ao Sistema de Banco de Dados Distribuído.

Independência de Rede de Comunicação : Topologia da rede, protocolo, método de acesso, etc, são tópicos transparentes ao sistema distribuído. O sistema de banco de dados distribuído deve rodar em qualquer plataforma de rede.

Independência de SGBD : A independência de SGBDs em um sistema homogêneo é muito fácil de ser alcançada. A interface de acesso padrão utilizada é a linguagem SQL, padronizada pela ANSI (American National Standards Institute). Os problemas se avolumam quando um banco de dados distribuído é implementado em um sistema heterogêneo de SGBDs. Atualmente existem comercialmente SGBDs hierárquicos, redes e relacionais. Rodar transações distribuídas envolvendo três formas distintas de armazenamento e de estruturação

dos dados é o grande problema que se apresenta. As soluções de intercomunicação entre SGBDs dão-se através de pontes (gateways), utilizando-se acesso nativo aos dados ou protocolos abertos (ODBC, DRDA, etc).

2.3.2 -Arquitetura de Referência para um Banco de Dados Distribuído

Nesta seção será apresentada uma arquitetura de referência que, embora não seja totalmente implementada neste trabalho, facilitará o entendimento da estrutura de um banco de dados distribuído.

Como é mostrado na fig. 1, um banco de dados distribuído, ao ser definido, deve possuir os seguintes componentes: Esquema Global, Esquema de Fragmentação e Esquema de Alocação [6].

No *Esquema Global* são descritas todas as entidades de dados (relações) e seus relacionamentos. A definição dos dados, em um banco de dados distribuído, é feita da mesma forma que em um banco de dados centralizado.

O *Esquema de Fragmentação* fornece o mapeamento entre relações globais e seus respectivos fragmentos. Determina como relações globais são divididas em fragmentos verticais, horizontais ou mistos. Desta forma, a conectividade do relacionamento entre relações globais e fragmentos é de 1:N.

O *Esquema de Alocação* define em que sítio cada fragmento irá residir. Determina de que forma os fragmentos são mapeados para alocações físicas.

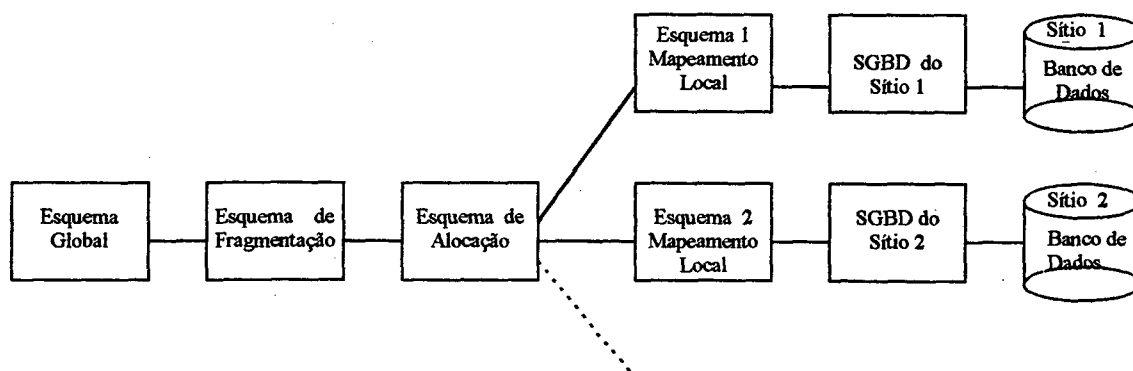


Fig. 1 - Representação da arquitetura de um Banco de dados distribuído.

2.3.3 - Projeto de Banco de Dados Distribuído

A arquitetura de um banco de dados distribuído é constituída de esquema conceitual global, esquema de fragmentação, esquema de alocação e o projeto físico do banco de dados em cada sítio. Na seção 2.3.2 descreveu-se o esquema global, esquema de fragmentação e esquema de alocação. No projeto físico mapeia-se o esquema conceitual global para áreas de armazenamento permanentes em cada sítio.

Para definir o esquema de fragmentação e alocação deve-se levar em consideração requisitos das aplicações sobre os dados; sítio onde a aplicação é usada; frequência de ativação; número, tipo e distribuição estatística dos acessos feitos para cada item de dado requisitado (tupla ou coluna).

O Projeto de Banco de Dados Distribuído deve ter como metas:

- *Processamento Local* : Deve-se reduzir acessos remotos e simplificar os controles dos programas de aplicação.

- *Disponibilidade e Confiabilidade dos Dados Distribuídos* : A utilização de cópias facilita o acesso em caso de indisponibilidade de um determinado sítio e aumenta a confiabilidade do sistema em caso de acidentes catastróficos na instalação.

- *Distribuição da Carga de Trabalho* : Aproveita a capacidade de cada sítio (armazenamento e processamento) e maximiza o nível de paralelismo de execução da aplicação.

- *Custo de Armazenamento e disponibilidade* : Deve ser analisado em cada sítio para se estudar alternativas de alocação e fragmentação.

Em um projeto de banco de dados distribuído deve-se ter como um dos principais objetivos a maximização do processamento local de cada sítio.

Existe basicamente duas propostas em um projeto de banco de dados distribuído :

- *Top-Down* : Geralmente utilizada em bancos de dados distribuídos homogêneos. Inicia-se pelo desenho do esquema global, seguido pelo esquema de fragmentação e alocação. Este modelo é complementado com o projeto físico dos dados alocados em cada sítio.

- *Bottom-Up* : Agrega bancos de dados existentes, integrando existentes esquemas de dados em um esquema global. Este modelo possui os seguintes requisitos : seleção de um modelo comum de banco de dados para escrever o esquema global; tradução de cada esquema local para o modelo comum de dados;

integração dos esquemas locais para o esquema global e resolução de conflitos entre os bancos de dados existentes .

Projeto de Fragmentação de Dados

O Projeto de Fragmentação de dados é o primeiro problema que deve ser resolvido no modelo Top_down de distribuição de dados. Este projeto consiste em identificar e agrupar tuplas (fragmentação horizontal) ou atributos (fragmentação vertical) que tenham a mesma propriedade do ponto de vista de sua alocação.

- *Fragmentação Horizontal* : É o particionamento das tuplas de uma relação global em subconjuntos disjuntos. A reconstituição da relação global se dá através da união de todos os fragmentos. Determinar a fragmentação horizontal significa determinar as propriedades lógicas dos dados (predicado de fragmentação) e propriedades estatísticas dos dados (números de referências das aplicações para o fragmento).

- *Fragmentação Horizontal Primária* : Determinar a fragmentação primária de uma relação global consiste na determinação de um conjunto disjunto e completo de predicados de seleção. O método para realizar esse tipo de fragmentação consiste de 2 passos :

1) Considera-se um predicado simples p_1 que particiona as tuplas de uma relação global R em 2 partes que são referenciadas diferentemente por pelo menos 1 aplicação. Seja $P = p_1$, onde P é um conjunto de predicados simples.

2) Considera-se um novo predicado p_i no qual particiona pelo menos um fragmento de P em 2 partes que são referenciadas em diferentes modos por pelo menos 1 aplicação. Seja $P \leftarrow P \cup p_i$. Elimina-se os predicados não relevantes de P . Repete-se este passo até que o conjunto de termos mínimos do fragmento P esteja completo.

- *Fragmentação Horizontal Derivada* : A fragmentação horizontal derivada de uma relação global R não é baseada em propriedades de seus próprios atributos, mas derivada da fragmentação horizontal de uma outra relação.

- *Fragmentação Vertical* : É o agrupamento de atributos de uma relação R . Cada atributo de R deve pertencer a pelo menos 1 fragmento e cada fragmento deve incluir um identificador único de tupla ou chave primária. A relação global pode ser obtida com a operação de junção dos diversos fragmentos. As alternativas para se chegar aos atributos de particionamento são:

a) Divisão : Consiste na divisão progressiva da relação global em fragmentos.

b) Agrupamento : Agregam-se progressivamente os atributos para se constituir os fragmentos.

- *Fragmentação Mista* : É o particionamento de uma relação global R utilizando-se a fragmentação horizontal e vertical.

A1	A2
	A3
	A4
	A5

Fig. 2 - *Fragmentação Vertical seguida por Fragmentação Horizontal.*

A4	A5	A1
		A2
		A3

Fig. 3 - *Fragmentação Horizontal seguida por Fragmentação Vertical.*

Exemplo de Fragmentação

Esquema Global

Emp(nr_matr,nm_emp,dt_nasc,nr_matr_ger,nr_depto)

Depto(nr_depto,nm_depto,nm_sigla,nr_matr_ger)

Forn(nr_forn,nm_forn,nm_sigla_cidade)

Prod(nr_prod,nr_forn,nr_depto,qt_prod)

Esquema de Fragmentação

Emp1 = SL depto <= 10 PJ

nr_matr,nm_emp,dt_nasc,nr_matr_ger,nr_depto (Emp)

Emp2 = SL 10 < nr_depto <= 20 PJ

nr_matr,nm_emp,dt_nasc,nr_matr_ger,nr_depto (Emp)

Emp3 = SL nr_depto > 20 PJ

nr_matr,nm_emp,dt_nasc,nr_matr_ger,nr_depto (Emp)

Emp4 = PJ

nr_matr,nm_emp,nr_depto (Emp)

Depto1 = SL nr_depto <= 10 (Depto)

Depto2 = SL 10 < nr_depto <= 20 (Depto)

Depto3 = SL nr_depto > 20 (Depto)

Forn1 = SL nm_sigla_cidade = "SC" (Forn)

Forn2 = SL nm_sigla_cidade = "SP" (Forn)

Prod1 = Prod SJ nr_forn = nr_forn (Forn)

Prod2 = Prod SJ nr_forn = nr_forn (Forn)

Projeto de Replicação de Dados

Em BDDs Relacionais, a realização de replicação de dados se dá através de cópias de dados feitas a partir de tabela(s) master(s). O objetivo principal do uso de réplicas é a melhoria de performance de acesso aos dados. Quando a cópia é somente de leitura recebe o nome de *extrato* ou *snapshot*. Quando uma cópia possui a característica de poder ser atualizada é denominada somente de réplica. As cópias podem ser completas ou parciais. As parciais utilizam o log da tabela master para efetuar as atualizações.

As cópias extrato podem assumir os seguintes formatos:

- Simples : Não possui mecanismo automático de atualização da cópia. São definidas a partir de comandos DDL ou DML.

-Timestamp : A renovação da imagem dos dados é efetuada a partir de mecanismo de tempo definido no catálogo.

- Refreshed :

* Periódico : Vinculada a períodos de tempo para a renovação parcial ou total dos dados.

* Imediato : Todas as cópias de uma tabela master são atualizadas no mesmo instante da atualização desta.

As réplicas possuem os seguintes tipos:

- Periódica : Num primeiro momento é feita uma cópia completa ao sítio destino. A partir daí as réplicas já podem sofrer atualizações. As atualizações efetuadas nas réplicas são periodicamente transferidas para a tabela master (ver fig. 4).

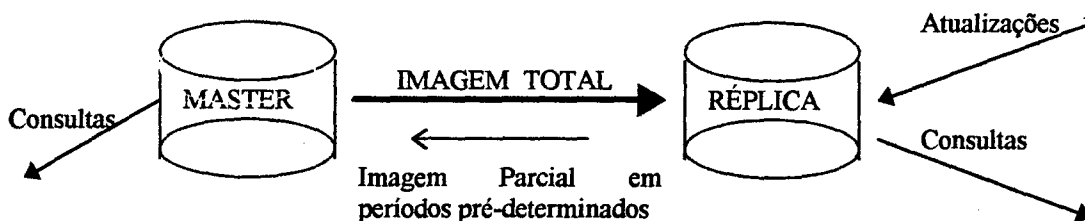


Fig.4 - Réplica Periódica

- Contínua : Após realizada uma cópia completa, esta passa a receber atualizações. Cada alteração realizada em uma réplica é imediatamente repassada à tabela master (ver fig. 5).

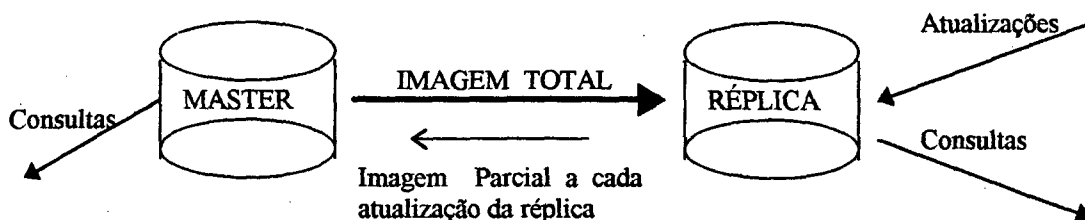


Fig. 5 - Réplica Contínua

No projeto de replicação de dados, incorporado pelo projeto físico do banco de dados distribuído, define-se para cada sítio as réplicas que comporão o banco de dados e a forma de atualização.

Projeto de Alocação de Dados

O Projeto de Alocação de dados tem por objetivo definir as alocações físicas dos arquivos do banco de dados distribuído em cada sítio, levando-se em consideração a máxima eficiência nas operações de consulta e atualizações dos dados.

O problema de otimizar a alocação dos arquivos nos sítios em um banco de dados distribuído tem sido arduamente estudado. Muitas das soluções são muito genéricas e complexas para serem usadas em problemas práticos, porque o número de parâmetros a ser considerado tende a atingir patamares proibitivos [4].

Em [4] é estudada a questão de como alocar os dados nos diversos sítios de um banco de dados distribuído de forma a tornar a consulta de uma transação o mais eficiente possível. São apresentadas as condições básicas para distribuição de dados nos aspectos qualitativo e quantitativo deste procedimento.

2.4 - Sistemas de Gerência de BDD's

2.4.1 - Arquitetura de um SGBDD

A arquitetura básica de um SGBDD é constituída pelo Diretório de Dados Global (DDG), Gerente de Transações (GT) e Gerente de Dados (GD) [5].

O DDG contém a descrição do banco de dados distribuído; o GT é o responsável pelo controle dos acessos ao BDD e o GD é a interface com o SGBD local, tendo como principal função as traduções entre banco de dados heterogêneos.

Em cada sítio residirá o SGBD local juntamente com uma cópia do SGBD global. Cada nó poderá ou não conter parte do diretório global. Isto dependerá da estratégia de alocação adotada.

2.4.2 - Gerência de Transações Distribuídas

A noção de transação é de fundamental importância no mecanismo de controle de concorrência. Uma transação é definida como uma série de ações, iniciada por um simples usuário/programa de aplicação, que deve ser tratada como uma unidade indivisível.

As propriedades básicas que devem ser satisfeitas por uma transação são [4] :

- 1) *Atomicidade* : Uma transação é uma unidade indivisível.
- 2) *Consistência* : Uma transação não altera a consistência de um banco de dados.
- 3) *Independência* : Uma transação executa independentemente de qualquer outra. Estados parciais de uma transação incompleta não são visíveis a outras transações.
- 4) *Durabilidade ou Persistência* : As ações realizadas pela transação, após o seu término, caso confirmadas, são permanentemente preservadas no banco de dados independente de falhas posteriores.

De uma forma geral, a Gerência de Transações tem como objetivo a eficiência, segurança e execução concorrente de transações.

Em um banco de dados distribuído, a gerência de transações tem como meta o controle da execução de transações, devendo garantir :

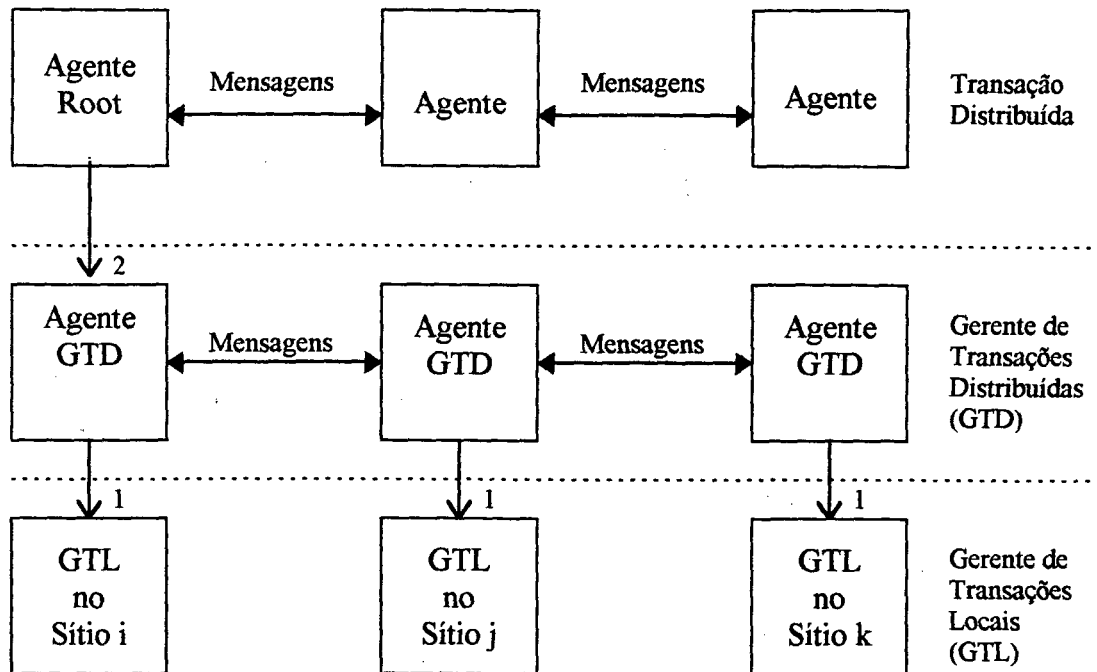
- 1) Atomicidade, durabilidade, seriabilidade e propriedades de isolamento.
- 2) Custos menores em termos de memória principal, CPU e número de mensagens de controle transmitidas e seu tipos de respostas.
- 3) Maximização da disponibilidade.

Recuperação de Transações Distribuídas

Para que haja atomicidade na execução de uma transação distribuída, duas condições são necessárias:

- 1) Em cada sítio participante, todas as ações são efetivadas ou não.
- 2) Cada sítio participante tem que ter a mesma decisão com respeito à confirmação ou cancelamento de sua subtransação.

O relacionamento entre a gerência de transações distribuídas e gerência de transações locais está representado o modelo de transações distribuídas apresentado na figura 6.



Interface 1 : Local_begin, Local_commit, Local_abort, Local_create

Interface 2 : Begin_Transaction, Commit, Abort, Create

Fig. 6 - Modelo Referencial de Recuperação de Transação Distribuída.

O Gerente de Transações Locais (GTL) implementa a interface 1: Local_begin, Local_commit e Local_abort. O GTL também pode criar processos agentes (Local_create).

O Gerente de Transações Distribuídas (GTD) é uma camada distribuída, implementada por um conjunto de agentes GTD locais, trocando mensagens entre eles. O GTD implementa a interface 2 : Begin_Transaction, Commit, Abort e Create agent (remoto).

No nível 3 temos as transações distribuídas, constituídas pelo agente root e outros agentes. Apenas o agente root pode usar as primitivas Begin_Transaction, Commit e Abort. A interface 2 é usada apenas pelo agente root.

O Gerente de Transações Distribuídas implementa a interface 2 assim:

Begin_Transaction : Quando o comando *Begin_Transaction* é usado pelo agente root, a GTD usará a primitiva *Local_begin* para a GTL do sítio de origem e para todos os sítios nos quais já existem agentes ativos da mesma aplicação, transformando todos os agentes em subtransações.

Abort : Quando o comando *Abort* é usado pelo agente root, todas as subtransações devem ser canceladas e recuperadas. Isto é feito usando-se o comando *Local_abort* para os GTLs de todos os sítios onde existe uma subtransação ativa.

Commit : Quando o comando *Commit* é usado pelo agente root, todas as subtransações devem ser confirmadas. Caso alguma subtransação fique impossibilitada de confirmar suas alterações, esta e todas as outras subtransações participantes devem recuperar as alterações realizadas no banco. Para este controle é utilizado o protocolo bifásico de confirmação (2PC).

2.4.3 - Gerência de Controle de Concorrência

O módulo responsável pelo controle de concorrência em um SGBD é conhecido como *Escalonador*. Sua principal função é evitar interferências entre transações. A comunicação entre o programa de aplicação e o *Escalonador* é realizada pelo *Gerente de Transações* (coordena operações de banco de dados na execução de aplicações). Qualquer técnica de controle de concorrência adotada pelo sistema deve garantir que toda execução concorrente de um conjunto de transações seja serializável. Isto equivale a dizer que, em um conjunto de transações, cada transação deve ser executada completamente antes que a próxima transação, envolvendo os mesmos dados, inicie.

Existem, basicamente, 3 técnicas de controle de concorrência : bloqueio, timestamp e método otimista.

Método de Bloqueio

O método de bloqueio é o método mais largamente utilizado. A regra básica consiste em bloquear um dado de forma compartilhada na leitura, e bloqueá-lo de forma exclusiva na escrita deste. A granularidade de um bloqueio diz respeito a sua abrangência: linha, página ou tabela como um todo.

O protocolo mais comum de bloqueio é conhecido como bloqueio de duas fases (2PL). Na primeira fase, conhecida como *fase de crescimento*, a transação

adquire os bloqueios e, na segunda, chamada de *fase de encolhimento*, os bloqueios adquiridos na primeira fase são liberados.

As características do protocolo 2PL são:

- 1) As transações são bem formadas, ou seja, uma transação que adquire o bloqueio em um dado antes de operá-lo e todos os bloqueios devem ser liberados após o término da transação;
- 2) As regras de compatibilidade de bloqueio são observadas, isto é, nenhum bloqueio conflitante é retido;
- 3) Uma vez liberado o bloqueio, nenhum novo é adquirido;
- 4) Todos os bloqueios de escrita são liberados juntos quando a transação é confirmada.

Quando se fala em bloqueio de recursos pode ocorrer um impasse conhecido como bloqueio mútuo (deadlock). O bloqueio mútuo ocorre quando uma transação espera por um recurso bloqueado por uma segunda transação, enquanto esta segunda transação também espera por um recurso bloqueado pela primeira transação.

O bloqueio mútuo em um sistema centralizado é geralmente detectado por um grafo *wait-for*. Em um grafo Wait-For as transações (no modelo centralizado) e agentes (no modelo distribuído) são representados por um nó e as requisições contendo bloqueios são representadas por setas. Como mostra a fig. 7, caso o grafo contenha um ciclo, existe bloqueio mútuo.

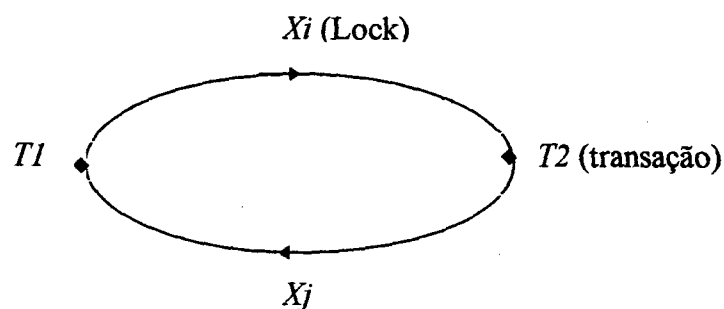


Fig. 7 - Grafo *wait-for* de detecção de bloqueio mútuo

A solução para resolver o bloqueio mútuo é abortar uma das transações. Para escolher a transação vítima, pode-se levar em consideração o volume de atualizações realizadas por cada uma das transações, tempo de execução, a causadora do bloqueio mútuo, etc.

A detecção de bloqueios mútuos em um ambiente distribuído é bem mais complexa, porque cadeias de transações, ou transações agentes, podem envolver um número diferente de sítios. Existem 3 propostas principais para se detectar o bloqueio mútuo em um Sistema Gerenciador de Banco de Dados Distribuído: centralizado, hierárquico e distribuído.

Centralizado : Todos os grafos locais *wait-for* são agrupados em um único sítio, conhecido como sítio de detecção de bloqueio mútuo. O grafo global é então examinado, verificando-se a ocorrência de ciclos. Problemas: sobrecarga na rede de comunicação e dependência de um sítio central.

Hierárquico : Os sítios na rede são organizados em uma hierarquia, como mostra a fig. 8. Um sítio local que possui um recurso bloqueado envia seu grafo *wait-for* para o sítio de detecção de bloqueio acima na hierarquia. Problema: difícil implementação, porém reduz o tráfego na rede.

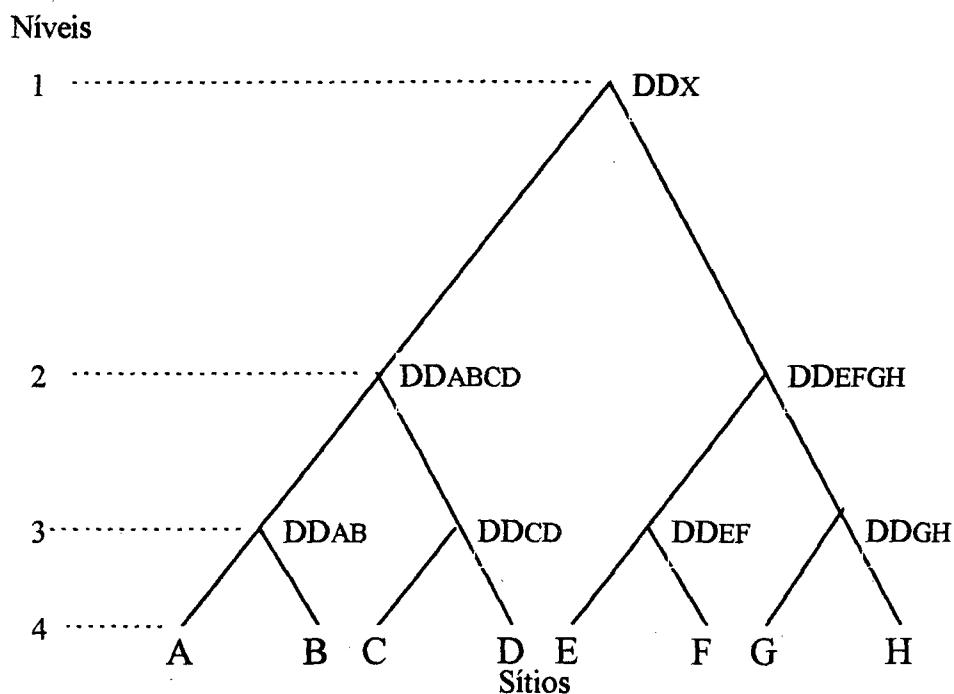


Fig. 8 - Árvore de detecção de Bloqueio Mútuo(Detector hierárquico) .

Distribuído : O método mais conhecido é o Obermarck. Neste método um nó adicional é introduzido - chamado EXT - em cada grafo *wait-for* local para indicar um agente (subtransação) no sítio remoto (ver fig. 9). Quando uma transação gera um agente em um outro sítio, um nó EXT é adicionado em ambos os grafos representando os sítios.

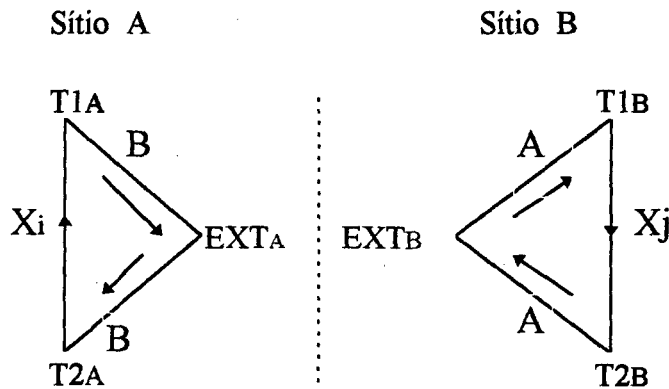


Fig. 9 - Detector de Bloqueio Global usando nós para representar agentes (EXT).

Para determinar se existe ou não bloqueio mútuo, é necessário juntar os dois grafos. Pode ser feito em qualquer dos dois sítios. O ciclo do bloqueio mútuo então aparecerá.

Método Timestamp

No método Timestamp não existe tempo de espera. Transações envolvidas em conflitos são simplesmente recuperadas e reexecutadas.

A regra fundamental do método Timestamp é ordenar as transações globalmente por tempo. Uma transação com menor timestamp tem prioridade na ocorrência de um conflito.

A leitura ou escrita de um dado por uma transação só é permitida se a última atualização do dado foi assinalada por uma transação anterior; caso contrário, a transação é reexecutada com um novo timestamp. Para cada transação é assinalado um timestamp.

Método Otimista

Como o próprio nome já diz, este método baseia-se no otimismo. Parte-se da premissa de que conflitos são raros e a melhor proposta é permitir que as transações processem sem controle de complexos métodos de sincronismo e sem qualquer tempo de espera. Quando uma transação confirma suas atualizações, o sistema verifica conflitos e, caso existam, a transação é reinicializada. A atomicidade da transação é garantida, porque as atualizações ocorrem em uma

cópia dos dados e são propagadas para o banco de dados somente quando nenhum conflito é detectado.

As transações são executadas obedecendo 3 etapas:

1) Fase de Leitura : Esta fase representa o corpo para a confirmação . Nada é escrito no banco de dados nesta fase.

2) Validação : As atualizações realizadas pela transação são examinadas quanto a conflitos.

3) Fase de Escrita : Se nenhum conflito foi detectado na fase de validação, então as atualizações são repassadas ao banco de dados. Se na segunda fase ocorreu erro de validação, a transação é desfeita e reinicializada.

2.4.4 - Gerência de Recuperação

A principal função da Gerência de Recuperação, em caso de falha do sistema, é identificar quais transações têm que ser desfeitas e quais têm que ser refeitas e, em seguida, efetuar as necessárias alterações no banco. O arquivo log, também chamado de journal ou audit trail, é a peça fundamental nesta função.

Arquivo Log

Um arquivo Log é um arquivo que registra todas as operações realizadas por todas as transações em um banco de dados.

Uma entrada é registrada no log cada vez que um dos seguintes comandos é emitido por uma transação : begin transaction, insert, delete, update, commit transaction, abort transaction, savepoint.

Para limitar a pesquisa em um log, em caso de falha, a gerência de recuperação insere marcas periódicas de sincronismo no log. Em caso de recuperação , o sistema retorna até a última marca.

Existem dois tipos de marca de sincronismo:

-Síncrona : O sistema não aceita novas transações até que todas as transações ativas tenham sido terminadas.

-Assíncrona : O sistema continua o processamento de transações quando a marca é registrada no log.

Existem basicamente 4 algoritmos usados nos protocolos de recuperação em um SGBD local:

- 1) Undo/Redo
- 2) Undo/No-Redo
- 3) No-Undo/Redo
- 4) No-Undo/No-Redo

Undo/Redo : É o mais complexo, pois envolve refazer e desfazer as transações após uma falha. Tem a vantagem de permitir ao gerenciador de buffer decidir quando descarregar os buffers, reduzindo assim a sobrecarga de I/Os. Produz a máxima eficiência durante a operação normal do sistema.

Undo/No-Redo : Os buffers do banco de dados são descarregados na operação de confirmação, não havendo necessidade de refazer transações na reinicialização do sistema, como também não há necessidade de gravar “after Images” no log.

No-Undo/Redo : O gerente de recuperação não escreve transações não confirmadas para o banco de dados. O gerenciador de buffer é forçado a reter os registros em memória principal no buffer de banco de dados até a confirmação de uma transação; as atualizações também podem ser escritas no log ao invés de escrevê-las no buffer de banco de dados.

No-Undo/No-Redo : Para evitar a operação de desfazer as atualizações das transações, o gerente de recuperação não escreve as atualizações para o banco de dados antes da confirmação de uma transação. Para evitar de refazer transações, o gerente de recuperação exige que todas as atualizações sejam escritas no banco. Este aparente paradoxo é resolvido escrevendo-se para o banco de dados em uma simples ação atômica na etapa de confirmação. As atualizações são escritas diretamente para o banco de dados via buffer. Endereços são gravados em uma lista de endereços, denominada de imagem. A única ação em uma confirmação é a atualização de índices de banco de dados, para apontar para a nova área, usando-se a lista de endereços imagem. Nenhuma ação é requerida na reinicialização do sistema.

Protocolo Bifásico de Confirmação

O Protocolo Bifásico de Confirmação é usado pelo gerente de transações para a conclusão de transações distribuídas em cooperação com o gerente de recuperação. A idéia básica do protocolo de confirmação bifásico (2PC) é determinar uma única decisão para todas as subtransações de uma transação global, com respeito a sua confirmação ou cancelamento, ou seja, garantir a atomicidade de uma transação distribuída.

O protocolo consiste de 2 fases. Na primeira fase é votada pelos nós envolvidos a decisão a ser efetiva. Na segunda fase, esta decisão é implementada.

Fase 1 : Na primeira fase o coordenador da transação (sítio onde a transação iniciou) pede aos sítios participantes para que se preparem para a confirmação. Antes de emitir o primeiro “prepare”, o coordenador grava no log um registro chamado “prepare log record”, no qual identifica todas as subtransações participantes. O coordenador também ativa o mecanismo de timeout, que o despertará quando um tempo pré-fixado expira.

O coordenador decidirá se irá confirmar ou cancelar a transação como um todo após receber todas as respostas dos sítios participantes. Se todos responderem com “ready”, ele decide confirmar a transação; caso contrário, a transação é cancelada.

Fase 2 : O coordenador inicia a segunda fase gravando sua decisão, que será “global_commit” ou “global_abort” no log; em seguida, ele informa a todos os sítios participantes sua decisão. Todos os sítios participantes escrevem um registro “commit” ou “abort” no log local baseado na mensagem recebida do coordenador. A partir deste momento, o procedimento de recuperação local é capaz de identificar que as ações da subtransação distribuída local não são perdidas.

Finalmente, todos os sítios participantes enviam uma mensagem final de reconhecimento “ack” para o coordenador e fazem a ação acordada. Quando o mesmo recebe a mensagem “ack” de todos os sítios participantes, ele grava um registro “complete” no log identificando o final da transação.

A seguir, são descritas as etapas básicas de um protocolo de confirmação bifásico:

Coordenador : Escreve registro “prepare” no log;
Envia mensagem “prepare” e ativa “timeout”;

Participante : Espera por mensagem “prepare”;
Se o participante está pronto para confirmar Então
Início
Escreve registro de subtransação no log;
Escreve registro “ready” no log;
Envia mensagem “ready” para o coordenador;

Fim
senão
Início
Escreve registro “abort” no log;
Envia mensagem “abort” para o coordenador;
fim

Coordenador : Espera por mensagem “ready” ou “abort” de todos os participantes ou por “timeout”;

Se expirou “timeout” ou alguma mensagem é “abort” Então
Início
Escreve registro “global_abort” no log;
Envia mensagem “abort” para todos os participantes;

Fim
Senão
Início
Escreve registro “global_commit” no log;
Envia mensagem “commit” para todos os participantes;

Fim

Participante : Espera por mensagem do coordenador;
Escreve registro “abort” ou “commit” no log;
Executa comandos locais;

Coordenador : Espera por mensagem “ack” de todos os participantes;
Escreve registro “complete” no log;

2.4.5 - Integridade e Segurança

Integridade

A integridade de um banco de dados está relacionada à consistência, exatidão, validade e precisão dos seus dados. A violação de integridade ocorre quando inserimos uma chave inválida em uma tabela, falha no algoritmo de controle de concorrência, falha na gerência de recuperação, etc.

As regras de integridade comumente utilizadas ao se definir uma tabela são:

- 1) Regra de Integridade de Domínio,
- 2) Regra de Integridade de Relação,
- 3) Regra de Integridade Referencial
- 4) Regras de Integridade explícitas, definidas utilizando-se gatilhos de banco de dados.

Atualmente, com as sofisticções dos SGBDs, tem-se meios de definir regras de integridade que reflitam no banco de dados as regras que regem os negócios de uma organização.

A solução de conflitos de regras de integridade local e global depende do nível de integração entre os vários sítios participantes em um banco de dados distribuído. Se existe um forte controle global, então as regras de integridade globais devem ter prioridades sobre as locais. Por outro lado, se a federação dos bancos de dados distribuídos não possuir praticamente nenhum controle global, as regras de integridade local devem prevalecer.

Ao se definir regras de integridade globais em um banco de dados distribuído, deve-se verificar as regras de integridade locais para que não ocorram conflitos.

Segurança

Cada SGBD local possui a responsabilidade da segurança de seus dados. A violação de segurança ocorre quando deliberadamente há uma tentativa de acesso não autorizado a determinados dados.

Identificação e Autenticação : Para se ter acesso aos dados de um determinado banco de dados, o usuário/programa de aplicação deve fornecer sua identificação seguida de sua senha. Para conexões remotas, pode-se usar o usuário ativo da presente sessão, ou informando-se usuário/senha.

Distribuição de Regras de Autorização de Acesso : Em um ambiente de banco de dados distribuído, as regras de autorização de acesso aos dados definidas para um sítio devem ser replicadas aos outros sítios participantes

Encriptação : A encriptação dos dados é uma forma de coibir acessos não autorizados. É também uma forma de resolver o problema de pessoas que burlam as regras de segurança do SGBD ou que tenham acesso aos frames do protocolo da rede de comunicação. A encriptação dos dados pode ser realizada através de vários métodos, entre eles os métodos Data Encryption Standart (DES) ou Public Key Cryptosystems (PKC).

Visão Global : Através de definições de visões, os dados são disponibilizados de uma forma restritiva, ou seja, limita-se o número de colunas acessadas como também o número de linhas em uma tabela. Através dos comandos Grant e Revoke poder-se-á limitar o acesso ou liberá-lo.

2.4.6 - Administração de Bancos de Dados Distribuídos

Em um sistema de bancos de dados distribuído homogêneo, o DBA exerce atividades tais como: otimização de alocação de fragmentos e replicação dos dados; monitora a performance do sistema e efetua correções; migra dados de um sítio para outro, entre outras. O Administrador de Banco de Dados (DBA) pode ser considerado como uma pessoa, ou grupo de pessoas, responsável pelo gerenciamento dos recursos dos dados dentro de uma organização

Em um sistema de bancos de dados distribuídos heterogêneo, em face da autonomia dos sítios, é o DBA local quem decreta quais dados devem contribuir para o banco de dados distribuído, se os dados serão somente de leitura ou leitura/escrita, etc.

Descrever as funções do DA e DBA em uma instalação torna-se bastante complicado. Cada empresa possui uma estrutura gerencial própria, algumas possuindo um profissional exercendo mais de uma função, outras improvisando profissional e função, e assim por diante. O Administrador de Dados (DA) é uma pessoa ou um grupo de pessoas cuja principal função é o gerenciamento lógico dos dados dentro de uma organização. A relação de atividades relacionadas abaixo, vinculadas ao DBA e DA, é apenas uma das várias formas de situar profissional/função. Algumas funções são específicas de bancos de dados distribuídos, outras (a maioria) são de uso geral (banco de dados centralizados, principalmente).

São funções do DA [9] :

- 01) Documentar os recursos globais de dados.
- 02) Estabelecer e acompanhar o uso de padrões de atributos, entidades, etc.
- 03) Ser um elo de ligação entre usuários globais, gerentes, pessoal de produção, administrador de rede e DBAs locais.
- 04) Resolver conflitos entre diferentes grupos de usuários/gerentes.
- 05) Definir regras globais de segurança e integridade.
- 06) Desenvolver e planejar os recursos de dados.
- 07) Definir o esquema global de dados.
- 08) Promover o compartilhamento de dados para evitar duplicação desnecessária de dados entre sítios locais e também aproveitar uma integridade completa.

São funções de um DBA [9] :

- 01) Projetar e implementar o Esquema Conceitual do banco de dados local (projeto lógico do banco de dados).
- 02) Projetar e implementar o esquema interno do banco de dados (desenho físico do banco de dados).
- 03) Projetar e implementar o esquema externo do banco de dados (visões locais).
- 04) Implementar regras de segurança local e regras de integridade que não estão no esquema local.
- 05) Monitorar performance e efetuar ajustes nos SGBDs locais.

- 06) Planejar e implementar reorganizações dos bancos de dados locais.
- 07) Documentar os bancos de dados e esquemas locais.
- 08) Definir as porções (fragmentos de dados) dos bancos de dados locais que irão contribuir em bancos de dados globais.
- 09) Mapear regras entre os fragmentos e os dados locais.
- 10) Estabelecer regras de conversão entre os dados locais e globais.
- 11) Especificar regras de segurança e integridade para os fragmentos, obedecendo as regras de nível global.

O Catálogo de Dados

O catálogo é formado por um conjunto de tabelas de metadados, isto é, tabelas de definições dos bancos de dados, dos usuários, das autorizações usuários/dados, etc.

Em um banco de dados distribuído, o catálogo contém a descrição das fragmentações, alocações dos dados e mapeamento para nomes locais.

Em um ambiente distribuído o catálogo pode ser encontrado nas seguintes formas:

- *Centralizado* : O catálogo reside em um único sítio. As principais desvantagens desta proposta são os congestionamentos de acesso e a dependência de um sítio central.

- *Totalmente Replicado* : O catálogo é completamente replicado em cada sítio. É uma proposta relativamente fácil de ser implementada, porém as atualizações do catálogo são um grave problema pelo fato de que devem ser propagadas imediatamente a todos os sítios. Este modelo pode acarretar um grave risco de inconsistência.

- *Distribuído* : O catálogo é fragmentado e replicado como qualquer relação em um banco de dados distribuído, ou melhor, é um banco de metadados distribuído. Os dados locais continuam a ser gravados normalmente na porção local do catálogo. As relações globais estariam definidas em porções replicadas ou fragmentadas do catálogo global.

Dicionário de Dados Global

O dicionário de dados é um repositório onde são documentados todos os aspectos de um determinado dado. O catálogo, neste contexto, pode ser considerado uma importante parte deste repositório.

No dicionário de dados encontram-se as seguintes informações:

- 01) Definições de dados, usuário e aplicações.
- 02) Informações para auxiliar o desenvolvimento de aplicações.
- 03) Especificação de regras de segurança e integridade.
- 04) Informações estatísticas de acessos aos dados e de performance.
- 05) Interface parcial para outros softwares.
- 06) Esquema de fragmentação e alocação, definições de visões globais.
- 07) Mapeamento de relações globais para locais.
- 08) Informações de redes.
- 09) Regras de segurança e integridade globais.
- 10) Dados estatísticos para otimização de consultas globais.

2.5 - Conclusão

O objetivo principal deste capítulo foi apresentar uma visão geral sobre Bancos de Dados Distribuídos e Sistemas de Gerência de Bancos de Dados Distribuídos.

Inicialmente apresentou-se os conceitos de Banco de Dados, Banco de Dados Distribuído, Sistema de Gerência de Banco de Dados e Sistema de Gerência de Banco de Dados Distribuído, com o intuito de deixar bem claro cada um dos assuntos abordados neste trabalho.

Na seção sobre Banco de Dados Distribuído, relacionou-se os requisitos básicos de um Banco de Dados Distribuído ideal, estabelecidos por Christian Date. O modelo implementado é analisado, no capítulo 7, quanto ao atendimento ou não desses requisitos. Viu-se também, de forma resumida, os principais componentes de um Projeto de Banco de Dados Distribuído : Projeto de Fragmentação, Projeto de Replicação e Projeto de Alocação de Dados.

O Sistema de Gerência de Banco de Dados Distribuído foi descrito principalmente abordando suas funções básicas : Gerência de Transação, Gerência de Concorrência e Gerência de Recuperação.

Este capítulo foi finalizado com a descrição das funções do DA e do DBA, envolvendo o desenvolvimento tradicional de sistemas e sistemas distribuídos.

Capítulo 3

SISTEMA CONTROLE DA FROTA DE VEÍCULOS

3.1 - Introdução

Neste capítulo descreve-se de forma gráfica e descritiva o sistema de informação Controle da Frota de Veículos da Celesc, que é o sistema de BDD implementado. A descrição é feita de forma resumida, para passar uma idéia geral de sua abrangência.

Através da análise do Diagrama do Fluxo dos Dados e do Diagrama de Entidade-Relacionamento é possível se ter uma idéia do porte do sistema de informação a ser desenvolvido de forma distribuída.

A relação dos aplicativos com os seus dados e a relação dos principais processos distribuídos são descritos com um pouco mais de detalhe, pois são utilizados nas análises finais deste trabalho.

As entidades com as respectivas descrições estão relacionadas no Anexo 1.

3.2 - Apresentação do Sistema

A função transporte é o terceiro maior custo da Celesc, sendo superado somente pelos itens compra de energia elétrica e despesas com pessoal, surgindo, assim, a necessidade de se ter um sistema moderno e eficiente que forneça dados que auxiliem na tomada de decisão.

Este sistema tem como objetivo disponibilizar às agências regionais todas as informações diárias referentes à sua frota de veículos, como também integrar as informações das regionais com as da administração central, formando um banco de dados de transporte da empresa.

Também é função do sistema apurar os custos a preço de mercado das funções de transporte, gerando e disponibilizando a toda empresa diversos indicadores gerenciais por veículo, características de veículo, atividade, órgão de vinculação e órgão usuário.

O objetivo maior da implantação deste sistema é a minimização dos custos operacionais e a otimização no uso da frota de veículos.

Este sistema poderia ser desenvolvido de forma centralizada, utilizando-se o SGBD DB2 no computador de grande porte, porém, este tipo de solução contraria a natureza distribuída do sistema e o limita em termos de funcionalidade, ferramentas, interface, disponibilidade e tempo de resposta.

Uma outra maneira de desenvolvimento que poderia ser adotada é a forma descentralizada. Esta solução atende em parte os requisitos do sistema (descentralização dos dados, disponibilidade, etc). Todavia, compromete a integração dos dados e torna muito complexa a implementação dos processos de natureza distribuída.

O desenvolvimento e a implantação de um sistema de banco de dados distribuído atenderão os requisitos de funcionalidade e acessibilidade das informações de forma transparente, com um tempo de resposta bastante superior em relação ao modelo centralizado, visto que grande parte dos acessos aos dados é feita localmente.

Espera-se que com a adoção deste modelo de desenvolvimento de sistemas na Celesc se esteja dando início a uma forma econômica e eficiente de gerenciar e manipular as informações, como também auxiliando a empresa no aspecto gerencial através de um modelo que prioriza a informação aonde ela realmente tem sua origem.

3.3 - Descrição do Sistema

O sistema será descrito basicamente em função de seus módulos componentes, levando-se em consideração a sua finalidade, composição e possibilidades de consultas.

Para se ter uma visão global são apresentados o Diagrama de Contexto, o Diagrama do Fluxo de Dados de nível 0, o Diagrama de Entidade-Relacionamento e o Diagrama do Fluxo de Informações.

3.3.1 - Diagrama de Contexto

No Diagrama de Contexto (Fig. 10) tem-se uma visão geral das áreas envolvidas na geração e uso dos dados do sistema. As principais áreas envolvidas são: Departamento de Recursos Humanos, fornecendo informações sobre os condutores; o Departamento Contábil, órgão responsável pela contabilização dos dados operacionais dos veículos e a Divisão de Transporte na Central e nas agências, que são as principais áreas usuárias. Este sistema também alimenta outras áreas da empresa e a Eletrobrás, através de dados estatísticos sobre o uso da frota de veículos.

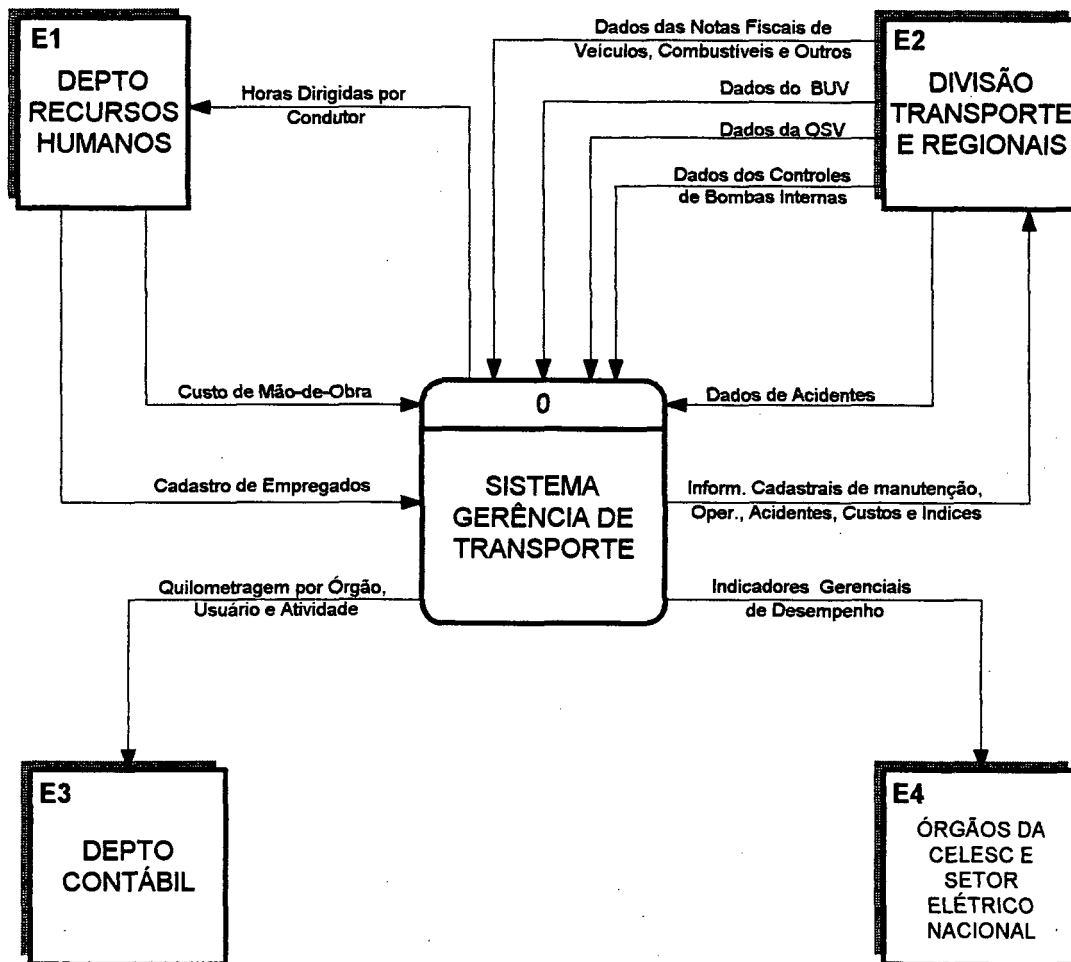


Fig. 10 - Diagrama de Contexto do sistema Controle da Frota de Veículos

3.3.2 - Diagrama de Fluxo de Dados de Nivel 0

O Diagrama de Fluxo de Dados de Nível 0 (Figs. 11 e 12) detalha o Diagrama de Contexto apresentado na figura 10. Pode-se ver neste gráfico os principais processos envolvidos, os depósitos de dados que dão origem às tabelas do sistema, e as áreas que geram e usam os dados.

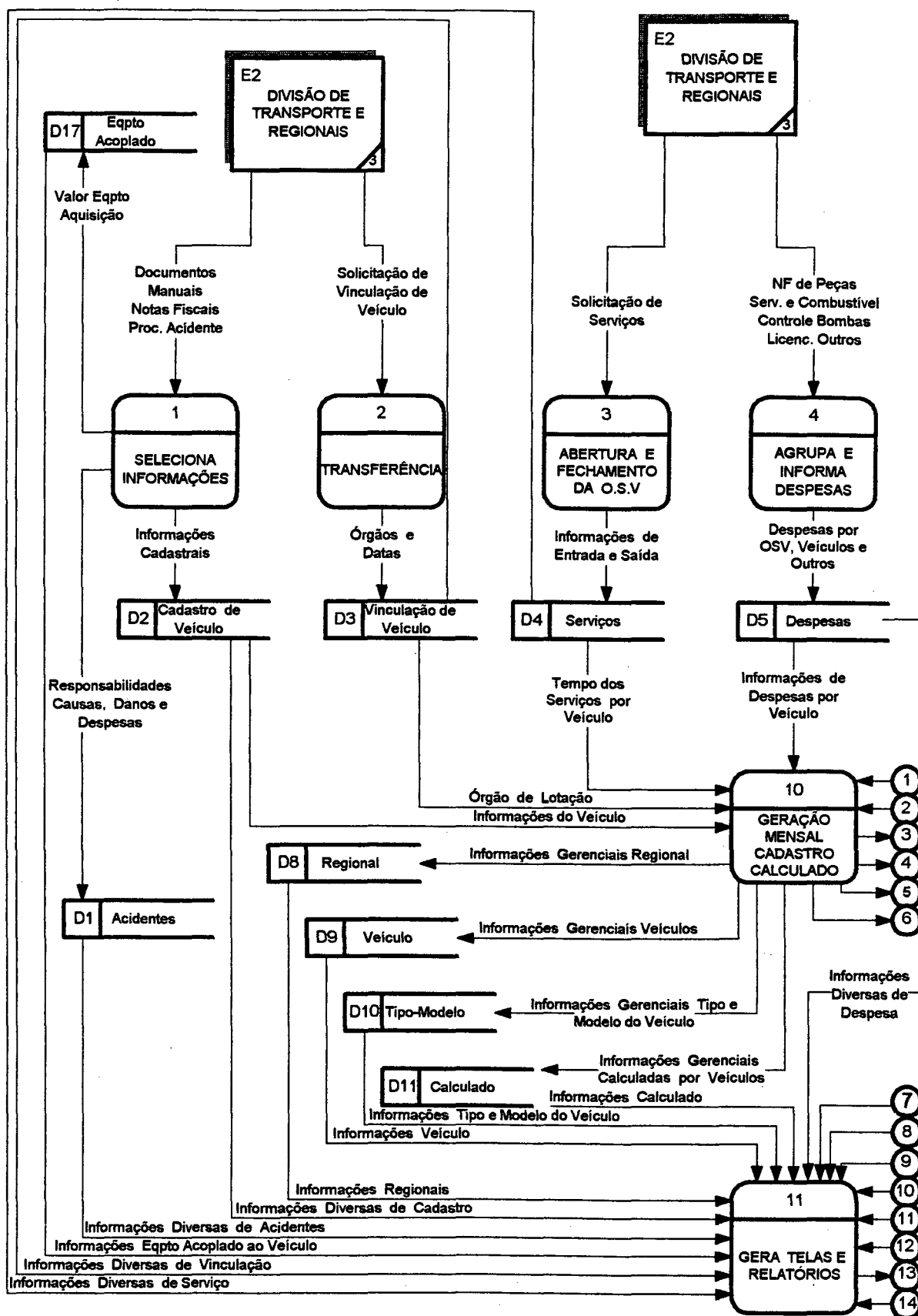


Fig.11 - Diagrama do Fluxo de Dados de nível 0 do sistema Controle da Frota de Veículos Parte 1

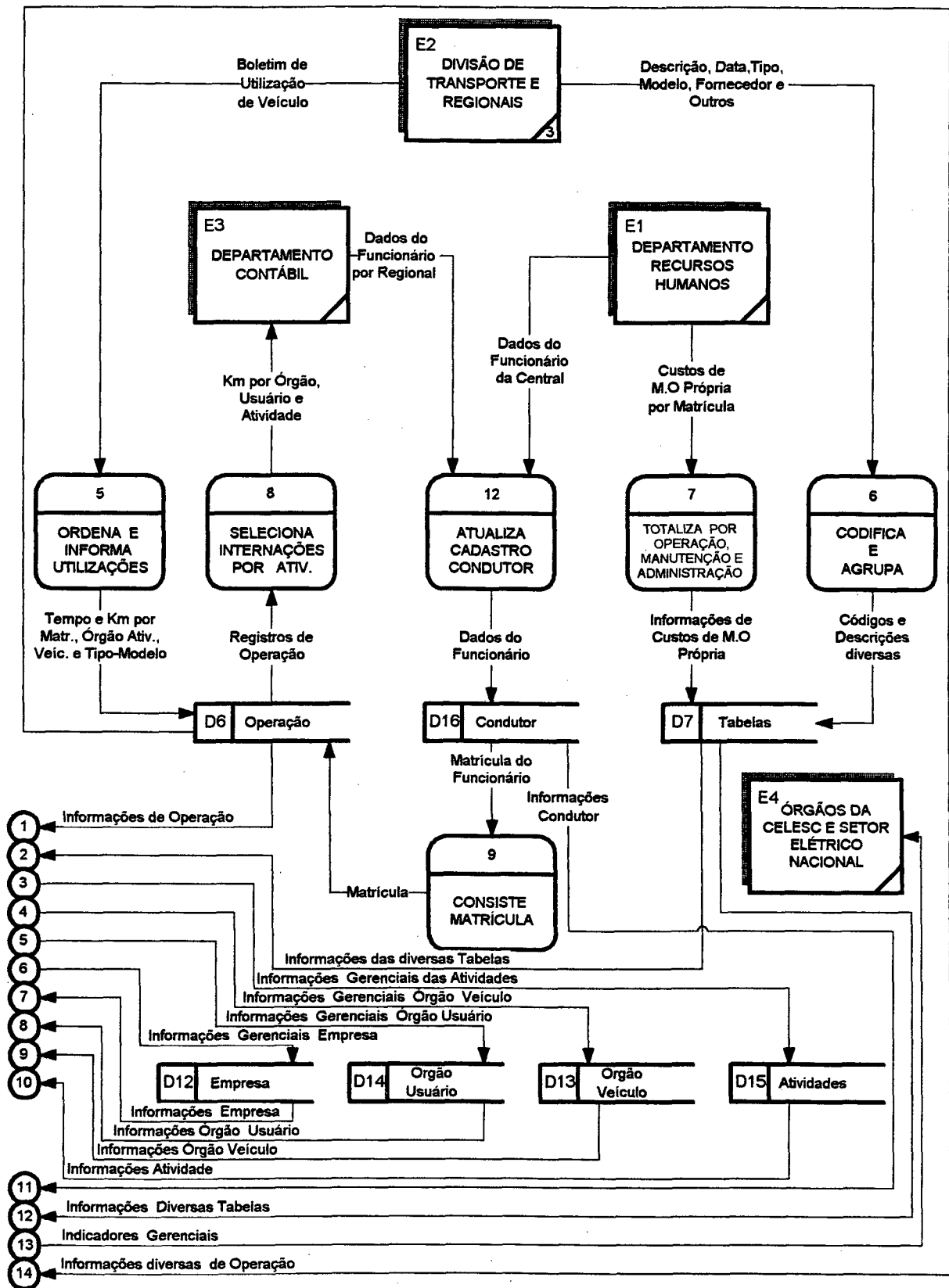


Fig. 12 - Diagrama do Fluxo de Dados de nível 0 do sistema Controle da Frota de Veículos Parte 2

3.3.3 - Diagrama Entidade-Relacionamento

No Diagrama de Entidade-Relacionamento (Fig. 13) apresenta-se as entidades do sistema com os seus respectivos relacionamentos. O modelo apresentado segue o método adotado pela Engenharia da Informação.

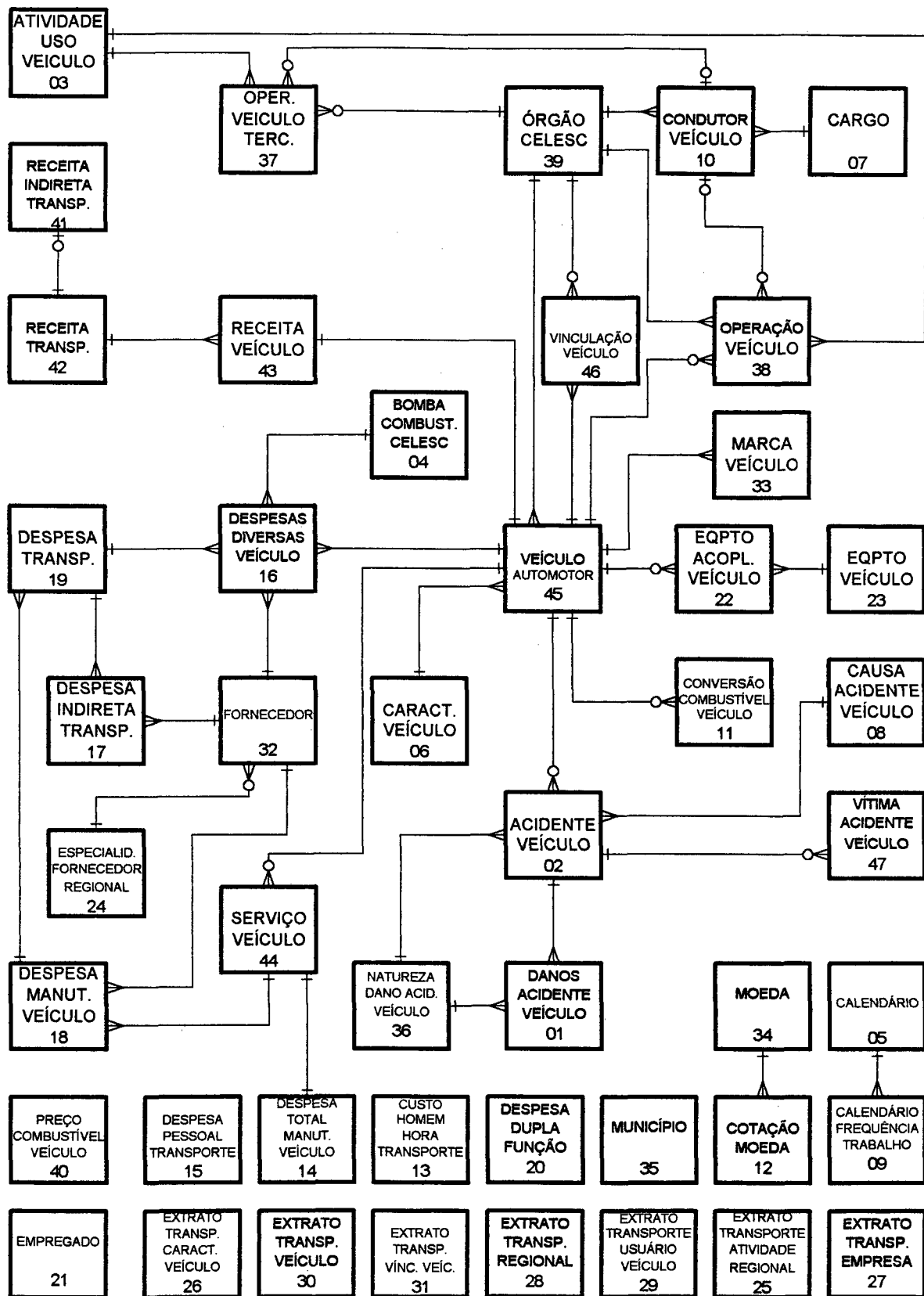


Fig. 13 - Diagrama de Entidade-Relacionamento do sistema Controle da Frota de Veículos

3.3.4 - Diagrama do Fluxo de Informações

Este diagrama mostra de forma resumida as atividades do sistema na agência central e nas agências regionais.

Como se pode observar no gráfico, grande parte do uso do sistema se dá nas agências regionais.

A agência Central é a responsável pela coleta de todos os dados mensais das agências para a geração dos índices gerenciais do uso da frota.

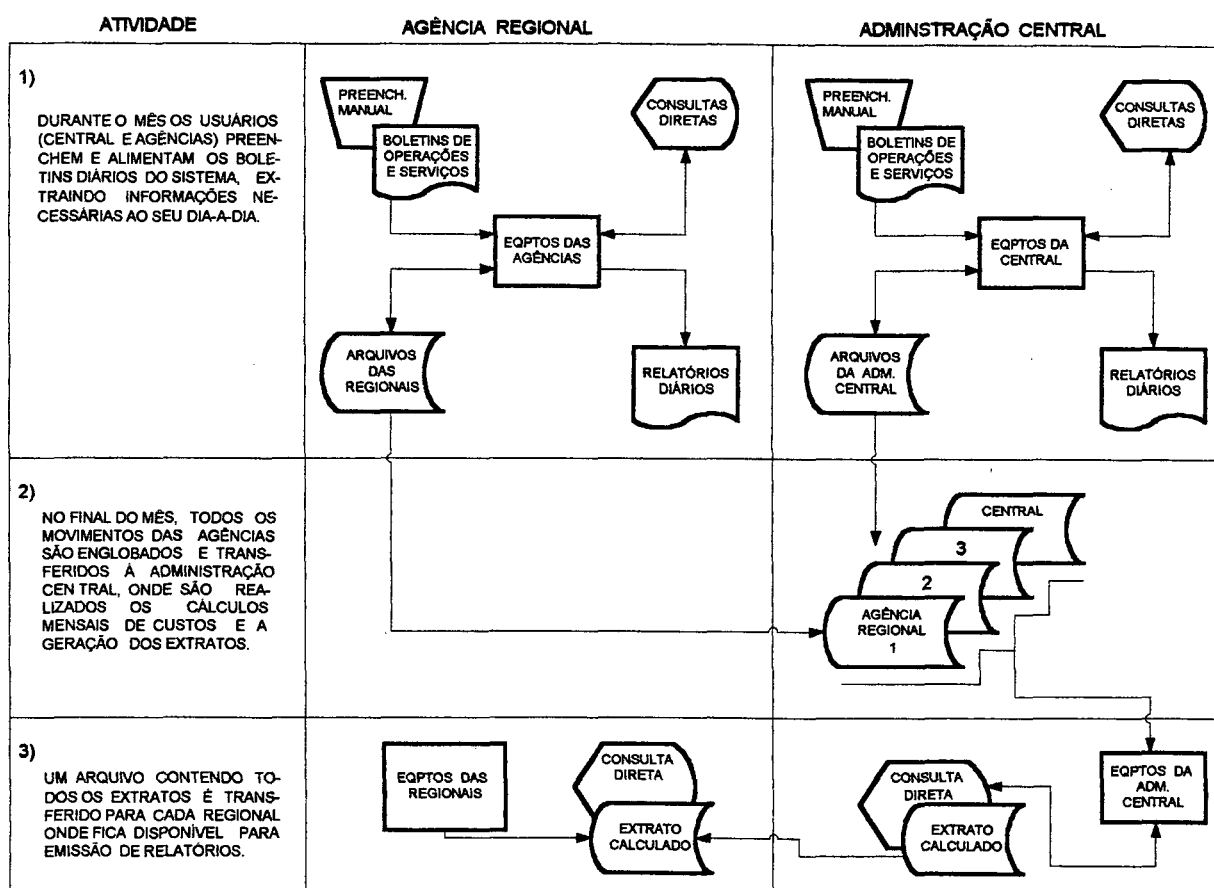


Fig. 14 - Diagrama do Fluxo de Informações do sistema Controle da Frota de Veículos

3.3.5 - Grade das Ocorrências Anuais das Principais Tabelas

Esta planilha detalha o volume de dados anuais inseridos no banco de dados de cada agência (são 16 no total).

Pelo volume dos dados apresentados pode-se ter uma idéia do porte médio deste sistema.

A atividade operacional dos veículos é o principal processo na geração de informações.

Tabelas	Central			Agências		
	Mínimo	Médio	Máximo	Mínimo	Médio	Máximo
CONVERSÃO-CMB-VEIC	20	50	100	200	500	1.000
DESP-PESSOAL-TRANP	10	20	30	200	400	600
VEÍCULO-AUTOMOTOR	200	250	300	1200	1.300	.500
FORNECEDOR-REGL	200	400	500	3.000	6.000	9.000
VINCULAÇÃO-VEÍCULO	200	250	300	1.200	2.000	3000
EQPTO-ACOPL-VEIC	1.000	1.200	1.500	10.000	13.000	15.000
OPERAÇÃO-VEÍCULO	30.000	50.000	100.000	600.000	900.000	1.500.000
OPER-VEIC-TERC	6.000	15.000	50.000	50.000	100.000	200.000
DESPESA-MANUT-VEIC	6.000	9.000	12.000	60.000	90.000	120.000
SERVIÇO-VEÍCULO	2.000	3.000	5.000	30.000	45.000	75.000
DESPESA-DVRS-VEIC	10.000	15.000	20.000	110.000	160.000	220.000
DESPESA-INDR-TRANP	10	20	30	100	200	300
RECEITA-VEÍCULO	100	150	300	1.000	1.500	3.000
RECEITA-INDR-TRANP	10	20	30	100	150	300
ACIDENTE-VEÍCULO	50	100	150	500	1.000	1.500
ACID-VEÍCULO-DANO	100	150	200	2.000	3.000	4.000
VÍTIMA-ACID-VEIC	30	50	100	200	400	600
EXTR-TRNP-CRT-VEIC	600	1.200	1.800	7.200	14.400	21.600
EXTR-TRNP-VEÍCULO	1.200	2.400	3.600	14.400	21.600	28.800
EXTR-TRNP-VINC-VEI	90	180	360	1.800	3.600	5.400
EXTR-TRNP-REGIONAL	12	18	24	240	360	480
EXTR-TRNP-USU-VEIC	150	300	450	2.700	3.600	5.400
EXTR-TRNP-ATVD-RGL	50	300	750	850	8.500	12.750

Fig. 15 - Planilha do volume dos dados do sistema Controle da Frota de Veículos

3.3.6 - Descrição dos Módulos do Sistema

Módulo 1 : *Cadastro de Veículo*

Finalidade : Nesse módulo estão cadastrados todos os veículos com os seus respectivos equipamentos e ferramentas. Nenhuma informação entra no sistema se o veículo não estiver devidamente cadastrado e, caso o veículo tenha sido retirado do cadastro, fica vedada a entrada de informações para o mesmo. Esse módulo é uma base de referência a todos os outros que integram o sistema.

Composição : Os dados que fazem parte desse módulo são : número de veículo, placa, marca, modelo, licenciamento, combustível, ano de fabricação, ano do modelo, cor, número do chassi, número do motor, quantidade de eixos, potência em HP, data da aquisição, valor de aquisição, valor complementar, valor atual, documento de aquisição, conta contábil, data do início da operação, órgão de vinculação, código livre trânsito, código de conservação, média km/l padrão, data da RDP (Retirada de Propriedade), número da RMD (Relação de Material Devolvido), número da Ordem de Desativação, código dos equipamentos acoplados, data da desativação, código dos equipamentos acoplados, data da aquisição dos equipamentos, valor dos equipamentos, código das ferramentas do veículo e quantidade de ferramentas.

Consultas : Esse conjunto de informações do cadastro permite uma série de relatórios dos veículos com suas características de acordo com as necessidades do usuário, como por exemplo: relação de veículos por órgão com as respectivas idades e estado de conservação, relação de veículos classificados por placa ou por tipo, entre outros.

Módulo 2 : Operação de Veículo

Finalidade : Nesse módulo são cadastrados diariamente todos os registros de operação dos veículos. Todos os deslocamentos dos veículos geram registros que entram no sistema através de digitação descentralizada, ficando cada órgão usuário responsável pela digitação das informações referentes aos veículos vinculados ao seu órgão. As consistências de todas as informações são realizadas na entrada dos dados, garantindo a confiabilidade das informações. Nesse módulo ficam armazenadas as informações referentes à quilometragem percorrida por

veículo, acessada diretamente pelo sistema contábil, e horas dirigidas por condutor, acessadas automaticamente pelo sistema de folha de pagamento.

Composição : Os dados que compõem esse módulo são: número do veículo, data da operação, local de saída, nome do usuário, órgão do usuário, matrícula do condutor, código de atividade, hora de saída e chegada, hodômetro de saída e chegada, local de chegada e conta contábil para algumas atividades.

Consultas : Esse conjunto de informações possibilita pesquisar todos os deslocamentos de um veículo ou de um condutor, entre muitas outras informações de fundamental importância para a empresa. A auditoria interna usa esse módulo diariamente para os trabalhos relacionados com a atividade transporte.

Módulo 3 : Manutenção de Veículo

Finalidade : Nesse módulo são cadastrados diariamente todos os registros de veículos gerados através das ordens de Serviço de Manutenção. São armazenadas todas as informações referentes aos custos de mão-de-obra própria, serviços de terceiros, peças, materiais e ainda o tempo do veículo em oficina.

Composição : Os dados que compõem esse módulo são : número do veículo, data, hora e hodômetro de entrada e saída do veículo em oficina, número da Ordem de Serviço, tipo de manutenção, tipo de despesa e valor.

Consultas : Esse conjunto de informações possibilita acompanhar os custos de manutenção do veículo, os componentes onde foram executadas as manutenções, o tempo do veículo na oficina, análise comparativa entre mão-de-obra própria e serviços de terceiros e ainda possibilita programar as manutenções preventivas.

Módulo 4 : Vinculação do Veículo

Finalidade : Esse módulo contém todos os veículos e seus respectivos órgãos de lotação. Toda vez que um veículo é transferido de lotação, um novo registro é incluído.

Composição : Os dados que compõem esse módulo são : número do veículo, data de vinculação, data final da vinculação e órgão.

Consultas : Esse módulo permite saber quais os veículos lotados em um determinado órgão, ou verificar em qual órgão está lotado um determinado veículo. Todos os relatórios do sistema por órgão necessitam acessar esse módulo.

Módulo 5 : Acidentes e Ocorrências

Finalidade : Nesse módulo são cadastrados todos os relatórios de acidentes e ocorrências com veículos e condutores da empresa.

Composição : Os dados que compõem esse módulo são : número do veículo, número do relatório do acidente, data do acidente, tipo do acidente, causa do acidente, danos com o veículo, despesas e danos ocorridos com as vítimas.

Consultas : Esse módulo permite saber quantos e quais veículos e condutores sofreram algum tipo de acidente, qual o custo da empresa com acidentes e quais os tipos de acidentes que são mais frequentes, além de outras consultas.

Módulo 6 : Credenciamento de Condutor

Finalidade : Esse módulo foi desenvolvido integrado com o módulo do banco de dados de pessoal, onde constam os dados cadastrais dos empregados. Cabe à Divisão de Transporte atualizar os dados referentes ao credenciamento de condutores para dirigirem veículos. Os condutores cadastrados somente são aceitos se estiverem devidamente credenciados a dirigir aquele tipo de veículo.

Composição : Os dados que compõem esse módulo são : matrícula do condutor, situação do condutor, cargo, data do credenciamento, categoria, data do descredenciamento, número da carteira de motorista, categoria, unidade da federação, data de validade, nome e telefone para contato em caso de acidente.

Consultas : Esse módulo é usado principalmente para consultas internas do sistema com o objetivo de consistir as informações de entrada dos dados de operação.

Módulo 7 : Controle de Combustível

Finalidade : Esse módulo surgiu de um desmembramento do módulo de Operação do Veículo, tendo em vista a adoção de uma nova política de controle de abastecimento e pagamento de combustível. Todo abastecimento interno ou via nota fiscal gera um registro nesse módulo, permitindo o acompanhamento do consumo de cada veículo.

Composição : Os dados que compõem esse módulo são : número do veículo, data do abastecimento, hodômetro do abastecimento e, para os abastecimentos nas bombas da empresa, o código da bomba e quantidade de combustível.

Consultas : Esse módulo permite acompanhar as médias dos veículos entre um abastecimento e outro, e também possibilita identificar possíveis irregularidades no consumo dos veículos.

Módulo 8 : Custos e Índices de Transportes

Finalidade : Esse módulo representa a síntese de todo o sistema de transporte, possibilitando ao gerente acompanhar o desempenho físico e econômico da frota. As telas e relatórios podem ser os mais variados possíveis, de acordo com o trabalho de cada gerente, a realidade de cada área e o interesse de cada agência.

Composição : Os dados que compõem esse módulo são : número do veículo, data de referência, tipo e modelo, órgão usuário, órgão de vinculação, código de atividade, mão-de-obra de manutenção, encargos de manutenção, serviços de terceiros, material de oficina, recapagem, lavagem, peças, óleo, pneus, combustível, depreciação, remuneração do capital, licenciamento, seguro, mão-de-obra de operação, encargos de operação, mão-de-obra de administração, encargos de administração, consumo de combustível, horas ofertadas, horas operadas, horas à disposição, horas de manutenção, horas de acidente, horas ociosas, índice de utilização, índice de disposição, índice de indisponibilidade, índice de avaria, índice de ociosidade, índice de km/l, índice de custo/km e km percorrido. Todos estes dados ficam armazenados num cadastro calculado por veículo, órgão, tipo, modelo, atividade, a nível geral da empresa. Todos os índices são gerados mensalmente.

Consultas : São disponibilizados para consultas todos os custos e índices de um determinado veículo, tipo e modelo, órgão de vinculação, órgão usuário e atividade. Os itens de custos gerados são agrupados em:

- Custo Mensal : Somatório do custo fixo mais custo variável.
- Custo Variável : Mão-de-obra de manutenção, serviço de terceiro, material de oficina, recapagem, lavagem, peças, óleo, pneus, combustível e diversos.
- Custo Fixo : Depreciação, remuneração do capital, licenciamento, seguro, mão-de-obra de operação, encargos de operação, mão-de-obra de administração e encargos de administração.
- Custo de Operação : Mão-de-obra e encargos de operação, licenciamento, seguro, recapagem, óleo, pneus e combustível.
- Custo de Manutenção : Mão-de-obra e encargos de manutenção, serviços de terceiros, material de oficina, lavagem, peças e diversos.

Os índices são gerados a partir do agrupamento das seguintes informações:

- Horas Ofertadas : São as horas referentes ao tempo útil mensal de um veículo.
- Horas Operadas : Somatório das horas em que os veículos estavam em deslocamento.
- Horas à Disposição : São os intervalos entre um deslocamento e outro num mesmo dia.
- Horas de Manutenção : Somatório das horas dos veículos em oficina.
- Horas de Acidente : Somatório das horas dos veículos em oficina devido a acidente.
- Horas Ociosas : É a diferença entre o total de horas ofertadas menos o somatório das horas operadas, à disposição e em manutenção.

A partir das informações de custos, quilômetros percorridos e agrupamento das horas envolvidas, tem-se os seguintes índices :

- Índice de Utilização : Relação entre as horas operadas, horas à disposição e horas ofertadas.
- Índice de Disposição : Relação das horas à disposição e o somatório das horas operadas, mais as horas à disposição.
- Índice de Indisponibilidade : Relação das horas de manutenção e horas ofertadas.
- Índice de Avaria : Relação das horas de acidente e horas de manutenção.
- Índice de Ociosidade : Relação das horas de ociosidade e horas ofertadas.

- Índice KM/Litro : Quilômetro percorrido em relação ao consumo de combustível.
- Índice KM/Hora : Quilômetro percorrido em relação às horas operadas.
- Índice Custo/KM : Custo total em relação aos quilômetros percorridos.
- Outros.

3.4 - Características Particulares (Processos Distribuídos)

A gerência da frota de veículos da Celesc possui características distribuídas que em um sistema descentralizado teria um certo grau de complexidade para equacionar algumas ações realizadas. A seguir são descritas algumas destas características:

Condutor : Cada agência possui somente os empregados vinculados a ela. Dados relativos a licenciamento, seguro e habilitação do condutor são realizados localmente. Outras informações cadastrais são realizadas via agência central e repassadas para as outras (16 no total) periodicamente.

Veículo : Cada agência possui somente os veículos vinculados a ela e poderá atualizar o cadastro de veículo nos seguintes casos:

- Transferência interna;
- Operações, tais como : quilometragem percorrida;
- Conversão do tipo de combustível;
- Licenciamento.

A agência central é responsável pelo cadastro geral de veículos, renovação da frota, transferência entre agências, empréstimo entre agências e retirada do veículo da frota.

Abastecimento : Nas agências de Blumenau, Joinville, Florianópolis e Lages, a Celesc possui bomba de combustível própria. Carros de outras agências podem, ao passar nestas localidades, efetuar abastecimento. O registro deste abastecimento é realizado na agência proprietária da bomba de combustível, porém deve ser transferido para a de vinculação do veículo.

Tabelas Básicas : As tabelas básicas do sistema residirão na RISC/6000 da agência central. Estas tabelas estão replicadas em todas as agências. Periodicamente ocorre a atualização destas cópias.

Extratos : No final de cada mês é disparado um processo em cada agência que calcula alguns indicadores e outros dados consolidados de despesas e operações. Este processo faz a transferência dessas informações para a RISC central onde são globalizadas a nível de empresa, atividade e característica de veículo.

Os extratos gerados na agência central são então replicados para cada agência, para consulta e tomada de decisões.

Empréstimo : Sazonalmente ocorrem empréstimos de veículos de uma agência para outra, principalmente em época de veraneio. O controle do veículo emprestado permanece na agência de origem, com suas operações e outras tabelas agregadas. Na agência de destino, ocorre o cadastro deste mesmo veículo, inicializando-se os serviços e operações.

No retorno deste veículo, seu cadastro permanece na agência que requisitou o empréstimo, com despesas e operações realizadas no período. Na agência de origem são inseridas as linhas de despesas e operações realizadas no período de empréstimo. Deve-se atribuir valor nulo ao atributo órgão e condutor.

O processo de empréstimo de veículo é gerenciado pela agência central.

Transferência : No processo de transferência, os dados do veículo permanecem na agência de origem com todas as suas tabelas agregadas (operações, despesas, receitas, etc). Na agência destino este veículo é cadastrado com todas as tabelas dependentes, colocando-se valor nulo nos campos condutor e órgão.

A agência central gerencia todo o processo de transferência.

3.5 - Principais Aplicativos

Os principais aplicativos do sistema de Controle de Frota são aqueles que envolvem o maior número de manipulações diárias de dados, ou seja, registro de

despesas e operações com veículos. Torna-se bastante difícil fazer estimativas de consultas, pois é um sistema inédito nas agências, e vislumbrar as possíveis consultas torna-se bastante difícil.

A seguir estão relacionadas as principais aplicações executadas diariamente pelo pessoal das agências, envolvendo os registros diários das atividades com veículos.

Despesas Diversas com Veículos : Aplicação responsável pelo registro de despesas diversas com veículo, tais como: combustível, licenciamento, seguro, etc.

<i>Tabelas Consultadas</i>	<i>Colunas(bytes)</i>	
Despesa_Transporte	ds_despesa_veículo(30)	
Fornecedor_regl	nr_form_fantasia(30)	
<i>Tabelas Atualizadas</i>	<i>Colunas(bytes)</i>	<i>Linhas/Dia</i>
Despesa_dvrs_veic	Todas(36)	40

Operações com Veículos : Registra a ocupação do carro com relação à atividade da solicitação de uso, tempo de utilização e quilometragem percorrida nesta atividade.

<i>Tabelas Consultadas</i>	<i>Colunas(bytes)</i>	
Veículo_automotor	le_hodm_fim_oper(06)	
<i>Tabelas Atualizadas</i>	<i>Colunas(bytes)</i>	<i>Linhas/Dia</i>
Operação_veículo	Todas(58)	195
Veículo_automotor	le_hodm_fim_oper(06)	195

Acidentes com Veículos : Informações referentes a um determinado acidente envolvendo veículo da empresa.

<i>Tabelas Consultadas</i>	<i>Colunas(bytes)</i>	
Condutor_veículo	nm_condutor_veic(30)	
<i>Tabelas Atualizadas</i>	<i>Colunas(bytes)</i>	<i>Linhas/Dia</i>
Acidente_veículo	Todas(76)	Indefinido
Veículo_automotor	id_situação_veic(02)	Indefinido

Serviços Realizados no Veículo : Registra todas as despesas gerais do veículo que envolvam uma parada significativa deste, por exemplo, troca de óleo, lavagem e lubrificação, manutenções, consertos, reparos, etc. Todas as atividades são realizadas via Ordem de Serviço de Veículo (OSV).

<i>Tabelas Atualizadas</i>	<i>Colunas(bytes)</i>	<i>Linhas/Dia</i>
Serviço_veículo	Todas(42)	04
Despesa_manut_veic	Todas(40)	10
Desp_manu_veic_tot	Todas(36)	04

Condutor do Veículo : Cadastro de funcionários da agência. É uma réplica particionada a partir do cadastro central de empregados da empresa. Mensalmente serão atualizadas a partir da agência central. Cada agência possui em média 330 funcionários, tendo cada linha deste cadastro 163 bytes. Na agência poderão ser alteradas 8 colunas desta tabela, totalizando 105 bytes. É uma tabela praticamente estática, sendo, porém, muito consultada.

Veículo : Cadastro dos veículos. É uma tabela particionada, controlada a partir da agência central. Alterações significativas ocorrem por ocasião de renovação de frota, incorporação de novos veículos ou leilões. Na agência podem ser alterados alguns dados cadastrais: licenciamento, seguro e vinculação interna.

<i>Tabelas Consultadas</i>	<i>Colunas(bytes)</i>	
Marca_veículo	ds_marca_veículo(20)	
Carac_veículo	ds_carac_veículo(30)	
Município	nm_município(20)	
Órgão_adm_celesc	nm_órgão_adm_ce(25)	
<i>Tabelas Atualizadas</i>	<i>Colunas(bytes)</i>	<i>Linhas/Dia</i>
Veículo_automotor	dt_ult_lccto_veic(8)	Indefinido
	dt_venc_seg_veic(8)	
	cd_mun_lccto_veic(3)	
	cd_org_vinc_veic(4)	

3.6 - Conclusão

A principal preocupação na descrição do sistema Gerência da Frota de Veículo foi dar uma visão geral dos seus dados e processos, detalhando-se os assuntos mais importantes para auxiliar nas análises da implantação distribuída.

A escolha desse sistema de informação teve como fatores determinantes o seu porte médio, favorecendo a sua implantação, as suas características distribuídas e a necessidade da Celesc no desenvolvimento descentralizado deste sistema.

A Celesc possui uma metodologia própria de desenvolvimento e documentação de sistemas. Adota o modelo conceitual, lógico e físico baseando-se na Engenharia da Informação. A metodologia da Celesc não contempla o modelo distribuído de um sistema de informação. As idéias desenvolvidas neste

trabalho serão apresentadas à equipe responsável pela metodologia na empresa para que esta passe a contemplar o desenvolvimento de sistemas distribuído.

Capítulo 4

DESCRIÇÃO DO AMBIENTE DE IMPLEMENTAÇÃO

4.1 - Introdução

Atualmente a Celesc desenvolve os sistemas corporativos (Contabilidade, Contas a Receber, Contas a Pagar, Materiais, Recursos Humanos, Orçamento, Distribuição e Geração de Energia, etc) de forma centralizada, utilizando um computador IBM 9121. As agências acessam as informações através de terminais da família IBM 3270, utilizando linhas Transdata de 4800 bps e 9600 bps.

Os sistemas com características descentralizadas são desenvolvidos na agência central e instalados em cada uma das agências, em equipamentos RISC/6000. As informações são acessadas através da arquitetura cliente/servidor, via rede e protocolo TCP-IP.

A adoção de banco de dados distribuído pela empresa exigirá algumas mudanças na metodologia de desenvolvimento de sistemas; novos conhecimentos para analistas, programadores e DBAs e, principalmente, melhoria na plataforma de comunicação de dados de longa distância.

Todo desenvolvimento e implementação do sistema distribuído de informação foi realizado utilizando-se as instalações da Celesc. Neste trabalho fez-se uso do ambiente de grande porte, do ambiente RISC/6000, rede local TCP-IP, rede de longa distância e microcomputadores da linha Intel.

Grande parte deste capítulo é dedicado ao SGBDD Oracle7, por ser a peça principal desta implementação. A replicação de dados e o protocolo bifásico de confirmação são as principais características a serem estudadas quando se fala de banco de dados distribuído utilizando-se o SGBDD Oracle7.

4.2 - Ambiente de Implementação

Para o desenvolvimento e implantação dos sistemas de informação centralizados, a Celesc conta hoje com um computador ES/9000 IBM 9121-732 com 3 processadores, 87 MIPs e 640 MB de memória; sistema operacional MVS/ESA; SGBD DB2 versão 3.0; como monitor de teleprocessamento o CICS versão 4.1; 2 impressoras laser IBM 3827 (90 ppm) e 1 impressora laser IBM 3160 (60ppm); 112 GB de disco rígido RAMAC; 3 controladoras de comunicações : 1 IBM 3725, 1 IBM 3745 e 1 IBM 3746-900, sendo que as duas últimas possuem funções de roteamento.

Para atender às necessidades locais das agências a Celesc desenvolve os sistemas em sua agência central, em equipamentos RISC/6000 e, após concluídos, instala-os em todas as agências regionais. A plataforma adotada é cliente/servidor, tendo-se como estação servidora uma RISC/6000 da IBM, de 64 MB e HD de 4GB, rodando o SGBDD Oracle7 versão 7.0.16, o software de comunicação SQLNET versão 1.0 e o protocolo TCP-IP. Nas estações cliente tem-se o Developer/2000 e o SQLNET V1 da Oracle, o protocolo TCP-IP da IBM, rodando sob Windows 3.11. Os microcomputadores são basicamente PENTIUM de 90 MHz, 32 MB de RAM e HD de 1.2 GB.

Na empresa adotou-se o DB2 da IBM como o SGBD responsável pelo gerenciamento dos bancos de dados corporativos da empresa. O Oracle7 é o SGBD destinado ao gerenciamento de banco de dados descentralizado e distribuído.

Este trabalho foi desenvolvido utilizando-se os recursos da Celesc. O SGBD Oracle7 foi adotado por ser o responsável pela implementação de banco de dados distribuído.

O acesso aos dados corporativos a partir do Oracle7 dar-se-á via gateway, denominado Transparent Gateway for DB2, da Oracle.

Neste trabalho fez-se uso de quatro sítios : um na agência de Blumenau, em Blumenau; um na agência de Florianópolis, no continente; e dois na agência central, na capital, sendo um de grande porte.

Para haver a comunicação entre SGBDs Oracle7 deve residir em cada sítio o software de comunicação SQL*NET, específico para um determinado protocolo na versão 1, e configurável para atender mais de um protocolo, na sua versão 2. Atualmente, na Celesc, utiliza-se o SQL*NET versão 1, para TCP-IP.

O desenvolvimento, implantação e testes do sistema de informação distribuído *Controle de Frotas de Veículos* da Celesc dar-se-ão através da plataforma Cliente/Servidor, utilizando-se como estação cliente um micro pentium de 90 mhz, 32 MB de RAM, 1.2 GB de disco rígido, rodando ferramentas de desenvolvimento Oracle sob Windows 95 e SQL*NET para TCP-IP. A figura 16 apresenta o ambiente básico do sistema de informação distribuído.

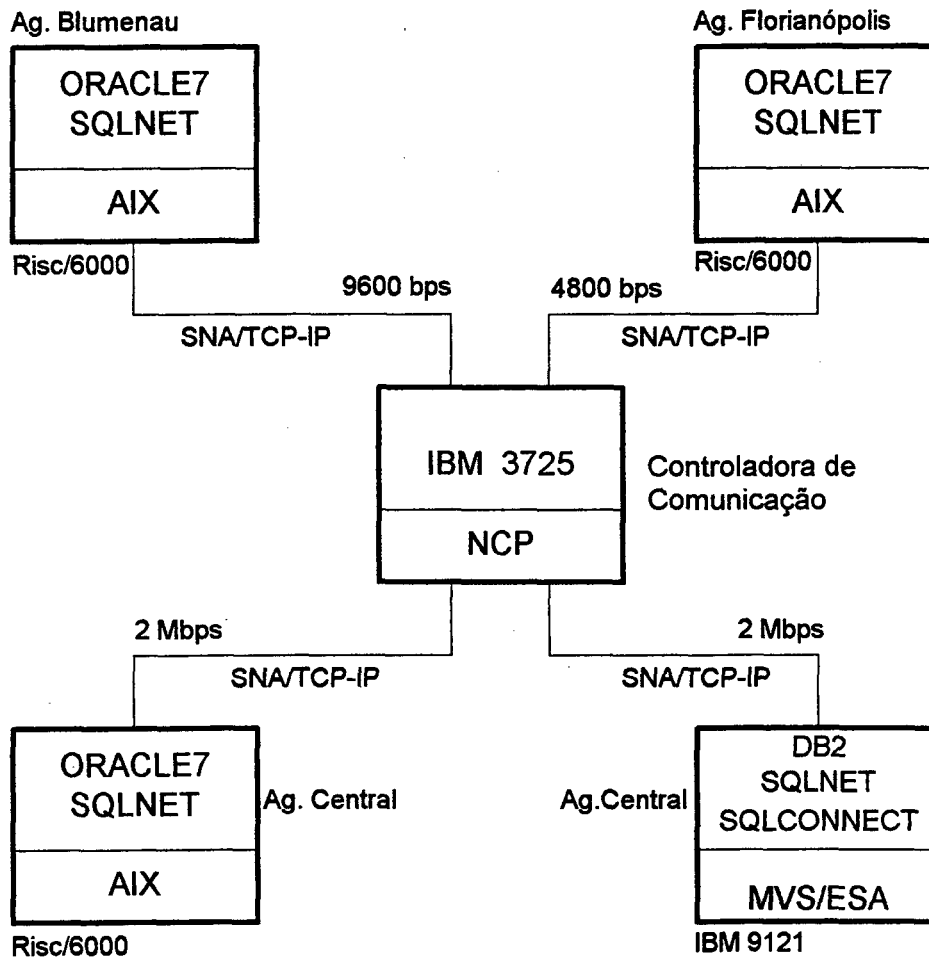


Fig. 16 - Topologia da rede de comunicação.

4.3 - Principais Características do SGBD Oracle7

O Oracle7 é o atual Sistema de Gerência de Banco de Dados Relacional da Oracle Corporation. Este Banco de Dados implementa o sistema de Banco de Dados Distribuído e foi utilizado para implementar o projeto piloto de banco de dados distribuído.

A seguir detalha-se as principais características deste SGBDD.

Controle de Concorrência

O Oracle7 emprega bloqueio a nível de linha para os dados e índices, propiciando assim um nível maior de concorrência. No SGBD Oracle7, a consulta não bloqueia a atualização, e o contrário também é verdadeiro, ou seja, a atualização de um determinado dado não evita o acesso concorrente ao mesmo. O Oracle7 possibilita consulta consistente, isto é, todos os dados de uma consulta são uma imagem do momento em que se iniciou a execução da mesma. Todas as atualizações (mesmo confirmadas) sofridas após ter sido iniciada uma consulta não são disponibilizadas a esta consulta.

O Oracle7 não emprega o bloqueio escalonável (linha, tabela, tablespace), ou seja, o bloqueio é sempre a nível de linha, a menos que a transação informe explicitamente o bloqueio em toda tabela.

Otimização de Consulta Baseada em Custos

O Otimizador do Oracle7 baseia-se em contabilizações estatísticas, tais como número de linhas da tabela, índices definidos, tamanhos médios de campos chaves, etc, para determinar o caminho de acesso mais eficiente para uma determinada consulta.

Periodicamente deve-se executar um utilitário que colete todas as informações estatísticas das tabelas dos bancos de dados, gravando-as nas tabelas do catálogo do SGBD. Ao se efetuar uma consulta, o otimizador irá acessar as informações estatísticas contidas no catálogo para criar o plano de acesso mais eficiente aos dados.

Limitador de Recursos

Com o Limitador de Recursos do Oracle7, o DBA consegue controlar os recursos do sistema que um determinado usuário pode possuir.

O DBA pode determinar o nível máximo dos recursos, tais como: tempo de CPU, I/Os lógicos em disco, tempo de conexão, etc. Se um usuário excede tais limites, sua sessão, ou requisição, é encerrada imediatamente.

Operações Contínuas

Funções cruciais tais como Cópia (Backup), Recuperação (Recovery) e operações com o SGBD podem ser realizadas de maneira “on line”, sem interromper o processamento de transações.

Padrão SQL

O Oracle7 é 100% compatível com o padrão SQL ANSI/ISO, propiciando um ambiente completo de desenvolvimento aberto de aplicações (16).

Restrições de Integridade

O padrão ANSI/ISO de restrição de integridade Referencial e de Entidade é completamente implementado pelo Oracle7.

As regras de integridade são definidas quando as tabelas são criadas ou alteradas e o SGBD Oracle automaticamente as ativa. O Oracle7 também implementa regras de domínio.

Procedimentos Armazenados e Gatilhos

Com Procedimentos Armazenados e Gatilhos PL/SQL (linguagem de programação Oracle), o Oracle7 possibilita o desenvolvimento de regras complexas de negócios a nível de Servidor. Estes procedimentos contêm comandos SQL e comandos da linguagem PL/SQL, e são armazenados no Banco em formato compilado de compartilhamento.

Os procedimentos podem ser chamados explicitamente via uma ferramenta de desenvolvimento Oracle (form, sqlplus, etc), como também ser executados automaticamente quando linhas são inseridas, atualizadas ou excluídas em uma tabela. Gatilhos podem ser utilizados para realizar regras complexas de integridade, para se fazer auditoria, para sincronizar tabelas replicadas, etc.

Segurança Baseada em Roles

Role é uma coleção de privilégios sobre uma ou mais tabelas e sobre outros objetos Oracle, agrupados e repassados a um usuário ou grupo de usuários. O DBA e usuários autorizados fazem a segurança das entidades de um determinado sistema via Role, Grant e Revoke.

Suporte a Línguas Estrangeiras (Nacional Language Suporte)

O Oracle7 fornece suporte a praticamente todas as línguas Asiáticas e Européias. Mensagens de erro, ordem de classificação e formatos de dados são alguns dos itens que são automaticamente convertidos para Linguagens Nativas.

Isto permite desenvolver aplicações cliente/servidor em plataformas heterogêneas e efetuar configurações em bases de dados distribuídas. A conversão de caracteres se dá de forma totalmente transparente.

Integração Transparente

O Oracle7 efetua integração transparente com grande parte dos SGBDs relacionais do mercado e também com outros não relacionais. Esta integração se dá através de gateways oracle, denominado Oracle Transparente Gateway.

Banco de Dados Distribuído

O mecanismo de replicação de dados e manutenção da integridade dos dados em transações distribuídas são os dois pontos fortes na implementação de bancos de dados distribuídos utilizando-se o SGBDD Oracle7.

A transparência de acesso aos dados é realizada através do uso de visões, sinônimos e gatilhos de banco de dados. Denomina-se gatilho de banco de dados o conjunto de comandos PL/SQL do Oracle7 que são executados antes ou após a inserção, atualização ou exclusão de uma determinada linha de uma tabela.

Processamento de Consulta Distribuída

Um simples comando SQL pode consultar dados de múltiplos Bancos de Dados e fazer de forma transparente consultas complexas.

O Usuário nunca necessita saber onde o dado está localizado fisicamente, e as aplicações nunca têm que ser recodificadas se um dado migra de um nó para outro.

Gerenciamento de Transações Distribuídas

O Oracle7 emprega um mecanismo transparente de Confirmação em Duas Fases (Two-Phase Commit) para implementar integridade de atualizações de transações distribuídas.

Transações que atualizam múltiplos nós são confirmadas (Committed) com o comando padrão SQL COMMIT. Nenhuma codificação se faz necessária. O Oracle detecta e resolve todas as falhas automaticamente, fazendo com que uma transação distribuída seja totalmente desfeita (Rollbacked) ou confirmada (Committed).

O Oracle7 Two-Phase Commit pode também interoperar com outros Gerenciadores de Banco de Dados, como o DB2, por exemplo.

Replicação de Tabelas

Dados comumente usados em múltiplos nós podem ser, transparentemente, replicados com o SGBDD.

O Oracle7 restaura automaticamente cópias, chamadas SNAPSHOT, a partir de uma tabela master, em intervalos de tempo definidos pelo usuário (horas, dias, semanas, etc).

A replicação pode se dar também somente com as linhas de uma tabela que sofreram alteração a partir do último SNAPSHOT.

Integração com Outros SGBDs

O Oracle7 consegue conversar com um grande número de outras fontes de dados através de seus produtos de conectividade SQL*CONNECT e Transparent Gateway.

Os dados podem ser acessados de vários sistemas:

- *Outros SGBDs Relacionais* : DB2, SQL/DS, RDB, NONSTOP SQL, etc.

- *Outros SGBDs não Relacionais* : IMS, HP TURBOIMAGE, IDMS, etc.

- *Sistemas de Arquivo Flat* : VSAM, RMS, ADABAS, etc.

- *Outras Fontes* : Machine Controllers, Stack Tickers, Process Monitor, etc.

Há também interoperação com um conjunto grande de linguagens de terceira geração (Cobol, C, Algol, etc) e produtos populares, tais como Lotus 1-2-3, Paradox, Foxpro, Excel, Word, além de outros.

4.4 - Mecanismos para Distribuição de Dados no Oracle7

Cada sítio, em um banco de dados distribuído Oracle7, necessita conhecer todos os outros sítios envolvidos. Este conhecimento se dá a partir da definição do "Database Link". Na definição de um banco de dados distribuído, pode-se utilizar Visões, Sinônimos, Procedimentos, Funções, Gatilhos e Snapshots.

Database Link

Define um caminho para um SGBD remoto. Pode ficar transparente ao usuário, utilizando-se, para tanto, de sinônimos e visões.

Ao se definir um *database link* deve-se especificar o seu tipo:

a) *Privado* : Possui vinculação a um determinado usuário. Pode ser referenciado apenas pelo proprietário.

b) *Público* : De domínio público.

c) *Rede* : Criado e gerenciado pelo serviço de nomeação de rede. Pode ser utilizado de qualquer SGBD da rede, especificando-se o nome global do objeto em um comando SQL ou definições no catálogo.

Visões

Podem ser usadas para providenciar transparência de localização para tabelas locais e remotas em um sistema de banco de dados distribuído.

Quando acessa uma visão, o usuário não sabe, ou não necessita saber, onde o dado está fisicamente gravado, ou se dados de uma ou mais tabelas estão sendo acessados.

Sinônimos

São muito úteis em ambientes distribuídos e não distribuídos, porque eles escondem a identidade do objeto, por exemplo, a localização deste. Se o nome de um objeto deve ser trocado ou movido, apenas o sinônimo deve ser redefinido; as aplicações baseadas em sinônimos continuam funcionando sem modificações.

Procedimentos Armazenados

A transparência de localização também pode ser efetivada através de unidades de programas PL/SQL contendo comandos SQL que referenciam dados remotos.

Gatilhos de Banco de Dados

O Oracle permite que se definam procedimentos que são implicitamente executados quando comandos SQL Insert, Update e Delete são usados em uma tabela. É similar a um Procedimento Armazenado, podendo conter comandos SQL e PL/SQL.

Na definição, pode-se especificar que o gatilho dispare para cada linha envolvida no comando SQL ou somente uma vez por comando. É também informado se o gatilho dispara antes ou após o comando envolvido.

Os usos mais comuns são:

- a) Geração automática de valores derivados de colunas;
- b) Prevenir transações inválidas;
- c) Implementar regras complexas de negócios;

d) Implementar integridade referencial através dos sítios em um banco de dados distribuído;

e) Implementar regras complexas de segurança;

f) Providenciar auditoria transparente de eventos;

g) Manter replicação síncrona de tabelas;

h) Gerar estatísticas de acessos às tabelas.

Um gatilho database possui três partes distintas:

a) Evento ou Comando do gatilho;

b) Restrição do gatilho;

c) Ações do gatilho.

Replicação

O Oracle possui dois mecanismos para transparentemente implementar réplicas de tabelas: *réplica síncrona* e *réplica assíncrona*. Réplica Síncrona é a cópia de uma ou mais tabelas cuja atualização se dá no mesmo instante em que a tabela master sofre alteração. Em uma Réplica Assíncrona, a atualização da cópia é postergada para períodos de tempos previamente definidos no catálogo.

As *réplicas síncronas* utilizam gatilho database e as *réplicas assíncronas* são definidas através do objeto *SnapShot*.

O uso de snapshots permite que uma determinada tabela seja replicada em um número ilimitado de sítios. Uma tabela snapshot pode ser somente de leitura ou pode também permitir atualizações. Tudo é controlado de forma automática e transparente, a partir da especificação do período em que for ocorrer as atualizações nas cópias ou tabela mestre.

A atualização de uma cópia pode ser feita de forma completa ou somente das linhas que sofreram atualizações. Para tanto deve existir tabela de log das alterações efetuadas.

O Protocolo Bifásico de Confirmação (2PC)

O Oracle automaticamente mantém a integridade dos dados envolvidos em uma transação distribuída através do protocolo de comprometimento de duas fases, denominado de *two-phase commit*.

É um mecanismo completamente transparente ao usuário, não exigindo nenhum esforço de programação. Este mecanismo, como o próprio nome já diz, possui duas fases distintas:

1 - *Preparação* (Prepare) : O Coordenador Global (sítio onde inicia a transação) pede aos nós participantes que se preparem para a fase de confirmação.

2 - *Confirmação* (Commit) : Se todos os nós participantes de uma transação distribuída respondem ao Coordenador Global que estão preparados, então este manda que todos se confirmem. Caso algum nó participante não possa se preparar, o coordenador global solicita que sejam desfeitas as alterações efetuadas pela transação em cada nó participante.

Em uma transação distribuída, no momento de execução é definida uma Árvore de Sessão, onde cada nó participante recebe uma denominação:

a) Cliente : Um sítio é denominado cliente quando faz referência a dados de um outro sítio.

b) Servidor : Um sítio é denominado servidor quando é diretamente referenciado em uma transação distribuída, ou é solicitado a participar em uma transação porque um outro sítio requisita dados dele.

c) Coordenador Global : É o sítio onde a transação distribuída se origina. Este nó torna-se a raiz da Árvore de Sessão.

d) Coordenador Local : É o sítio que referencia dados em outros sítios para completar sua parte na transação distribuída.

e) Ponto de Confirmação : O mecanismo de confirmação de duas fases sempre designa um nó *ponto de confirmação* na árvore de sessão. O trabalho deste nó consiste em obter uma ação atômica que determine se a transação será confirmada ou desfeita. Este nó distingue-se dos demais devido a dois usos:

- Nunca entra em estado de preparação. Sua função é gravar os dados críticos de uma transação.

- O resultado da transação neste nó determina como a transação será completada.

Cada SGBD recebe um "Commit Point Strength" na entrada do SGBD. O SGBD com o maior valor deste parâmetro, ao participar em uma transação distribuída, ficará sendo o sítio ponto de confirmação.

Na figura 17 é mostrada a atuação do mecanismo de confirmação de duas fases envolvendo três nós distribuídos, executados de forma bem sucedida.

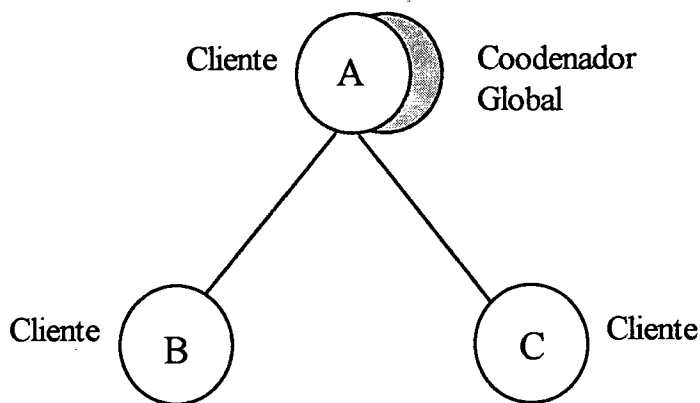


Fig. 17 - Protocolo de Confirmação de Duas fases envolvendo 3 sítios.

- 01 - A aplicação solicita um "COMMIT" da transação distribuída;
- 02 - O Coordenador Global "A" diz ao servidor "B" para se preparar;
- 03 - O servidor "B" se prepara;
- 04 - O servidor "B" informa ao Coordenador Global "A" que está preparado;
- 05 - O Coordenador Global "A" informa ao servidor "C" para se preparar;
- 06 - O servidor "C" se prepara;
- 07 - O servidor "C" informa ao Coordenador Global "A" que está preparado;
- 08 - O Coordenador Global "A" confirma-se;
- 09 - O Coordenador Global "A" informa ao servidor "B" para se confirmar;
- 10 - O servidor "B" confirma-se;
- 11 - O servidor "B" informa ao Coordenador Global "A" que se auto-confirmou;
- 12 - O Coordenador Global "A" informa ao servidor "C" para se confirmar;
- 13 - O servidor "C" confirma-se;
- 14 - O servidor "C" informa ao Coordenador Global "A" que se auto-confirmou;
- 15 - O Coordenador Global "A" dá por encerrada a transação;
- 16 - Os Servidores "B" e "C" também.

4.5 - Conclusão

O ambiente de implementação apresentado é uma visão parcial do que realmente foi a implantação do sistema distribuído Controle da Frota de Veículos da Celesc. Para o seu pleno funcionamento, urge que se instale o banco de dados e os aplicativos em cada uma das 16 agências regionais. O gateway com o DB2 deverá estar em funcionamento e, em cada agência, deverá estar disponível a rede local trafegando TCP-IP para que os aplicativos acessem os dados do banco de dados distribuído na forma cliente/servidor.

Os recursos de informática da Celesc estão sendo redimensionados para atender à demanda crescente de informações e, para 97, já se está planejando o aumento da capacidade das RISCs/6000 das agências e atualização do sistema operacional AIX e SGBDD Oracle7 para versões mais recentes. O SQLNET, software responsável pela comunicação entre SGBDs Oracle7 e entre aplicativos, também será atualizado para a versão 2.

Toda estrutura de comunicação de dados está sendo padronizada para se usar o protocolo TCP-IP, nas redes locais e de longa distância, nos microcomputadores, nas RISC/6000 e no computador IBM de grande porte. As linhas de comunicação sofrerão uma melhora significativa a partir do próximo ano, concluindo-se todo o projeto de melhoria no ano de 1998. Hoje todas as linhas das agências vêm até a agência central nas velocidades de 9600 e 4800 bps. Pela nova estrutura, chegarão na agência Central linhas de 8 Mbps provenientes de algumas agências concentradoras com funções de roteamento, eliminando assim a atual configuração em estrela única para estrela distribuída. Até o ano 2000 pretende-se ter todo o estado ligado por fibra ótica, utilizando-se os cabos para-raio das torres de transmissão da Eletrosul e Celesc, podendo-se chegar a uma velocidade de 2 Gbps.

Caso se tenha realmente concretizado todos os atuais planejamentos de melhoria das instalações, o banco de dados distribuído terá um grande alento para que se torne uma maneira prática e eficiente de desenvolver sistemas de informação na empresa.

O Projeto físico do banco de dados distribuído que hoje terá que sofrer algumas modificações para se ajustar a algumas limitações de comunicação e software, num futuro bem próximo poderá ser implantado de acordo com o seu desenho lógico.

Capítulo 5

PROJETO LÓGICO DISTRIBUÍDO

5.1 - Introdução

O projeto lógico distribuído tem como principal objetivo indicar em cada sítio do banco de dados distribuído a forma de distribuição (particionado, centralizado, replicado e extrato) a nível de macro assuntos (subsistemas).

Para o desenvolvimento deste capítulo, fez-se uso de uma metodologia que trabalha os processos do sistema (oriundos da análise dos processos), as entidades (levantadas a partir do modelo conceitual) e os locais físicos de abrangência do sistema distribuído [9].

A metodologia é composta basicamente de 4 etapas, cada qual definindo uma matriz específica. O principal objetivo é a geração da matriz Plano de Compartilhamento Lógico, que servirá como insumo do projeto físico. O Plano de Compartilhamento Lógico descreve os subsistemas componentes do sistema e a forma de distribuição dos dados.

As matrizes formadas nesta etapa são as seguintes:

a) Processos x Local Físico - Esta matriz relaciona os processos do sistema e os locais de origem destes.

b) Processos x Assuntos de Informação (Subsistemas) - Relaciona os processos do sistema com cada um dos subsistemas, informando o uso ou não do processo pelo subsistema.

c) Assunto de Informação x Local - Determina os locais físicos de cada subsistema e a natureza de usabilidade das informações dos subsistemas (cria, usa e não utiliza).

d) Plano de Compartilhamento Lógico - Esta matriz informa a maneira de armazenamento (replicada, particionada, centralizada e extrato) de cada assunto de informação em cada sítio do banco de dados distribuído.

A partir do levantamento de dados e processos junto às áreas usuárias do sistema de informação Gerência da Frota de Veículos, chegou-se aos seguintes assuntos de informação :

- a) Veículo
- b) Operação com Veículos
- c) Acidentes com Veículos
- d) Receita com Veículos
- e) Despesas com Veículos
- f) Informações Acessórias
- g) Dados Estatísticos.

Os locais físicos de abrangência do sistema são as agências regionais e a administração central.

5.2 - Matriz Processos x Locais

Nesta etapa do projeto lógico distribuído é construída uma matriz que relaciona os principais processos do sistema e os locais físicos de abrangência.

O intuito é verificar em cada sítio pertencente ao banco de dados distribuído o uso de um determinado processo.

PROCESSOS

Processa Receitas							
Processa Operações							
Processa Despesas							
Processa Acidentes							
Mantém Veículos							
Mantém Tabelas Acessórias							
Gera Informações Sumarizadas							
LOCAL							
Administração Central	X	X	X				
Agências				X	X	X	X

5.3 - Matriz Processos x Assuntos de Informação

Uma matriz Processos x Assuntos de Informação relaciona os assuntos de informação (subsistemas) e os principais processos. Cada relacionamento indica a natureza deste, ou seja, a forma como cada assunto de informação é utilizado em cada processo (não utiliza, utiliza e cria).

ASSUNTOS DE INFORMACÃO

Veículo
Operações com Veículos
Acidentes com Veículos
Receitas com Veículos
Despesas com Veículos
Posicionamento das Agências no Uso da Frota
Tabelas Acessórias

PROCESSOS

Processa Receitas	U			C			U
Processa Operações	U					C	U
Processa Despesas	U	C					U
Processa Acidentes	U				C		U
Mantém Veículos	U						C
Mantém Tabelas Acessórias	C						
Gera Informações Sumarizadas	U	C	U	U	U	U	U

C= Cria U = Utiliza BRANCO = Não Utiliza

5.4 - Matriz Assuntos de Informação x Locais

A matriz Assuntos de Informação x Locais relaciona os assuntos de informação com os locais de abrangência do sistema. Em cada local é informada a condição de uso de cada assunto de informação (não utiliza, utiliza e cria).

ASSUNTOS DE INFORMAÇÃO

Veículos								
Operações com Veículos								
Acidentes com Veículos								
Receitas com Veículos								
Despesas com Veículos								
Posicionamento das Agências no Uso da Frota								
Tabelas Acessórias								
LOCAL								
Administração Central	C	C						C
Agências	U	U	C	C	C	C	C	U

C = Cria U = Utiliza BRANCO = Não Utiliza

5.5 - Plano de Compartilhamento Lógico

O Plano de Compartilhamento Lógico informa, para cada local de abrangência do sistema, a maneira como cada assunto de informação será armazenado.

Antes de montar a matriz de compartilhamento lógico, passar-se-á a elencar alguns motivos de centralização e distribuição de dados e formas de distribuí-los.

Quando Centralizar?

- Dados utilizados por aplicações centralizados.
- Usuários de várias áreas necessitam da posição atual destes dados, que são freqüentemente atualizados.
- Usuários movem-se por diversas localidades utilizando-se sempre os mesmos dados.
- Dados como um todo sofrem muitas consultas, como por exemplo, dados gerenciais.
- Necessidade de alto nível de segurança dos dados.
- Volume de dados muito grande, não havendo espaço suficiente para armazenamento, a não ser no banco de dados central.

Quando Distribuir?

- Dados utilizados no local, raramente ou nunca, são utilizados por outros locais.
- Segurança, integridade e privacidade dos dados é de responsabilidade local.
- Dados locais são pesquisados ou manipulados através de linguagens de usuário final.
- Necessidade de alta performance no acesso aos dados.
- Custos altos na transmissão de dados.
- Disponibilidade de dados é importante a nível local.

Formas de distribuição

Particionado(Fragmentado) :

- A maioria das transações origina-se e é processada onde se localiza.
- Alto volume de processamento local.
- Necessidade de alto nível de performance.
- Dados de outros locais são raramente acessados e não há processamento sobre os dados como um todo.
- O assunto de informação é de natureza fragmentada,ou seja, são subconjuntos disjuntos de dados.

Centralizado :

- Quando não atende os requisitos de fragmentação, o assunto de informação deve residir em um ambiente centralizado.

Replicado :

- Necessidade de se evitar transmissão de dados.
- Relação armazenamento/custo de transmissão é favorável.
- Possibilita ganho de disponibilidade.
- Frequência de acessos de consultas é bem superior à frequência de atualização.
- Melhoria de performance.
- Podem sofrer atualizações.

Extratos :

- Muito utilizados por sistemas de apoio à decisão, onde os dados não necessitam estar atualizados sincronamente.
- Formatação de dados de maneira a atender melhor as consultas locais.
- Dados não sofrem atualizações.

ASSUNTOS DE INFORMAÇÃO

Veículos								
Operações com Veículos								
Acidentes com Veículos								
Receitas com Veículos								
Despesas com Veículos								
Posicionamento das Agências no Uso da Frota								
Tabelas Acessórias								
LOCAL								
Administração Central	C	C						R
Agências	E	E	P	P	P	P	P	P

E = Extrato **R** = Réplica **C** = Centralizado **P** = Particionado

5.6 - Conclusão

O projeto lógico do banco de dados do sistema de gerência da frota de veículos teve como resultado final a relação dos subsistemas e sua forma de distribuição nos locais de abrangência do sistema. Esta matriz servirá como insumo da próxima etapa deste trabalho.

Para esta etapa utilizou-se partes resumidas de uma metodologia que envolve os 3 níveis definidos pela ANSI para gerenciamento da informação : conceitual, lógico e físico [9].

O aspecto importante desta etapa do desenvolvimento do sistema é se ter o modelo conceitual representando com fidelidade o mundo real, a essência do negócio, independente da tecnologia a ser adotada, ou de qualquer restrição existente atualmente na empresa. No nível físico, deve-se adaptar o modelo lógico de acordo com as características da tecnologia disponível.

Capítulo 6

PROJETO FÍSICO DISTRIBUÍDO

6.1 - Introdução

Neste capítulo será descrito o modelo físico de implementação do sistema distribuído Controle da Frota de Veículos. Serão apresentadas somente partes essenciais do projeto físico, porque se fosse aplicada a metodologia [9] em sua totalidade, o trabalho ficaria muito extenso e não traria muita contribuição ao mesmo.

No projeto lógico do sistema foi visto como cada subsistema será armazenado em cada local físico de abrangência, descrito através do Plano de Compartilhamento Lógico.

Nesta etapa do desenvolvimento do sistema será estabelecido o local físico de cada tabela. Um Assunto de Informação (subsistema) visto no projeto lógico poderá dar origem a um ou mais Assuntos de Distribuição. O Assunto de Informação, como já descrito no capítulo 5, é a divisão do sistema em subsistemas, gerando um agrupamento de entidades (tabelas) de acordo com a função que o subsistema desempenha dentro do sistema. Geralmente um Assunto de Informação dá origem a um Assunto de Distribuição. Porém, quando se detalha um subsistema e se observa detalhes de implementação física, um Assunto de Informação pode gerar mais de um Assunto de Distribuição. Neste trabalho um Assunto de Informação dá origem a somente um Assunto de Distribuição.

Este capítulo é composto dos seguintes assuntos:

a) Assuntos de Distribuição : Relaciona os subsistemas, obtidos a partir do detalhamento do Plano de Compartilhamento Lógico.

b) Identificação dos Locais de Distribuição : Relaciona os locais físicos de abrangência do sistema.

c) Definição da Forma de Distribuição : Especifica para cada assunto de distribuição a forma de atualização, ou seja, identifica se será síncrona ou assíncrona.

d) Distribuição Física dos Assuntos de Distribuição : Relaciona todos os assuntos de distribuição e suas tabelas componentes e os locais físicos de armazenamento, especificando a forma de armazenamento (centralizada, local, extrato, réplica e particionada).

e) Especificação da Distribuição : Descreve, para cada tabela, detalhes de seu armazenamento e particularidades de distribuição.

6.2 - Assuntos de Distribuição

Para se chegar aos assuntos de distribuição, utiliza-se como insumo o Plano de Compartilhamento Lógico, definido no projeto lógico, juntamente com o Diagrama de Entidade-Relacionamento. Aplica-se, a partir daí, o método da Categoria de Associação de James Martin [13].

A *Categoria de Associação* mede o grau de intensidade (variando de 1 a 7) com que uma entidade está associada a outra em um relacionamento. Quando o grau de intensidade do relacionamento é acima de 4, significa dizer que as tabelas devem permanecer juntas.

(1) - Quase nunca.

(2) - Raramente.

(3) - Ocasionalmente.

(4) - Regularmente.

(5) - Frequentemente.

(6) - Quase sempre.

(7) - Sempre.

Exemplo : Empréstimo Bancário.

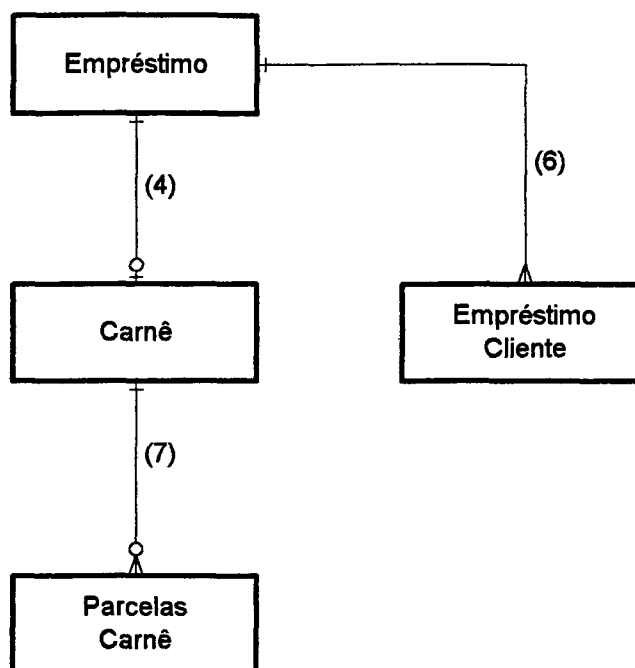


Fig. 18 - Exemplo do uso da Categoria de Associação.

No sistema distribuído Controle da Frota de Veículos, os Assuntos de Informação são os próprios Assuntos de Informação, ou seja, cada Assunto de Informação deu origem a somente um Assunto de Distribuição. São eles:

- a) Veículo.
- b) Operações com Veículos.
- c) Despesas com Veículos.
- d) Receitas com Veículos.
- e) Acidentes com Veículos.
- f) Tabelas Acessórias.
- g) Índices de Uso de Veículos.

Cada um dos assuntos de distribuição deveria estar representado pelo seu respectivo diagrama E-R. Para efeito de exemplificação será apresentado somente o diagrama do assunto de distribuição Acidentes com Veículos.

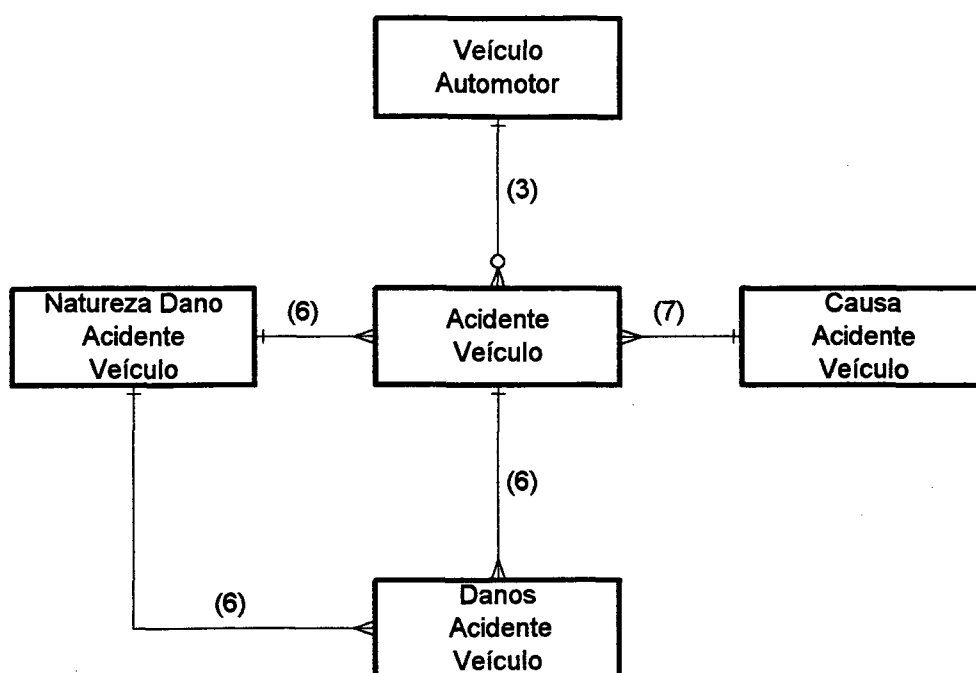


Fig. 19 - Assunto de Distribuição Acidentes com Veículos.

6.3 - Definição dos Locais de Distribuição

O local físico visto no Projeto Lógico pode desmembrar-se em mais de um local no Projeto Físico. O modelo físico gera novos detalhes que não são vistos a nível lógico. Neste trabalho, o local físico Agências do projeto lógico poderia dar origem a locais distintos como Gerência de transporte, Divisão de Transporte, Gerência Regional, etc. Como todos estes locais possuem as mesmas características funcionais e manipulam as mesmas informações, no projeto físico eles são representados pelo local físico Agências. No projeto físico deste trabalho se terá somente dois locais físicos representativos : a agência central e as agências regionais.

No projeto físico, os detalhes de implementação, a tecnologia atual disponível na empresa, a metodologia de desenvolvimento adotada e os recursos de hardwares e softwares determinam algumas adaptações que podem modificar bastante o projeto lógico.

6.5 - Distribuição Física dos Assuntos de Distribuição

O plano de compartilhamento lógico culmina com o plano de compartilhamento físico de dados. Esta matriz apresenta os assuntos de distribuição em cada um dos locais de abrangência do sistema, mostrando a forma de atualização de cada um deles.

Cada assunto de distribuição deveria ser detalhado a nível de tabelas para formar o plano de compartilhamento físico de dados. A título de exemplificação será mostrado o assunto de distribuição Veículo.

ASSUNTO DE DISTRIBUIÇÃO

Veículo-Automotor							
Característica-Veículo							
Conversão-Combustível-Veículo							
Equipamento-Veículo							
Equipamento-Acoplado-Veículo							
Marca-Veículo							
Vinculação-Veículo							
LOCAL							
Administração Central		C		C		C	C
Agências	L	E	L	E	L	E	P

C = Centralizado L = Local E = Extrato R = Réplica P = Particionado

6.6 - Especificação de Distribuição

Para cada tabela de cada assunto de distribuição deve ser feita a especificação de distribuição. Descreve-se a identificação da tabela segundo a metodologia adotada na empresa, seu nome por extenso, o local da tabela master, se for extrato ou réplica, o local de armazenamento, a forma de distribuição e particularidades de distribuição.

Esta etapa será exemplificada com a tabela Veículo-Automotor, do assunto de distribuição Veículo.

<i>Identificação</i>	<i>Nome</i>
Veículo-Automotor	Veículo Automotor da Empresa

<i>Local do Master</i> Administração Central

<i>Local</i> Agência	<i>Forma de Distribuição</i> Fragmentada/Replicada
-------------------------	---

<i>Particularidade de Distribuição</i>
<ul style="list-style-type: none"> - O particionamento se dará pelo número da agência do Veículo. - Chave dos fragmentos : Número da Agência e Número do Veículo - Algumas alterações de atributos (licenciamento, seguro) nas agências devem ser repassada sincronamente à tabela centralizada.

6.7 - Plano de Compartilhamento Físico

Nesta etapa relaciona-se, para cada local do projeto físico, as tabelas com a respectiva forma de distribuição.

A partir desta descrição tem-se o mapeamento de todas as tabelas do sistema distribuído, agrupadas pela forma de distribuição.

Para que o modelo de compartilhamento físico de dados fosse mais representativo, ter-se-ia que considerar também aspectos de rede, acesso aos dados e recursos da instalação da empresa. Porém, o método apresentado possibilita um dos caminhos para a distribuição dos dados.

Agências

Réplicas

- Empregado
- Calendário
- Calendário_Frequência_Trabalho

Extratos

- Atividade_Uso_Veículo
- Marca_veículo
- Equipamento_Veículo
- Causa_Acidente_Veículo
- Característica_Veículo
- Despesas_Transporte
- Serviço_Veículo
- Natureza_Dano_Acidente_Veículo
- Moeda
- Cotação_Moeda
- Despesa_Dupla_Função

- Custo_Homem_Hora_Transporte
- Município
- Cargo
- Extrato_Transporte_Característica_Veículo
- Extrato_Transporte_Vinculação_Veículo
- Extrato_Transporte_Regional
- Extrato_Transporte_Usuário_Veículo
- Extrato_Transporte_Atividade_Regional
- Extrato_Transporte_Empresa
- Extrato_Transporte_Empresa
- Extrato_Transporte_Veículo

Tabelas Locais

- | | |
|--------------------------------|---------------------------------|
| - Operação_Veículo | - Despesa_Indireta_Transporte |
| - Órgão_Celesc | - Especialidade_Fornecedor_Reg |
| - Condutor_Veículo | - Acidente_Veículo |
| - Operação_Veículo | - Vítima_Acidente_Veículo |
| - Vinculação_Veículo | - Danos_Acidente_Veículo |
| - Receita_Veículo | - Serviço_Veículo |
| - Receita_Transporte | - Despesa_Manutenção_Veículo |
| - Bomba_Combustível | - Despesa_Pessoal_Transporte |
| - Equipamento_Acoplado_Veículo | - Preço_Combustível_Veículo |
| - Despesas_Diversas_Veículo | - Receita_Indireta_Tranporte |
| - Despesa_Transporte | - Conversão_Combustível_Veículo |
| - Fornecedor | |

Fragmentos

- Veículo_Automotor (agência) =

$$SL\ NR_AGÊNCIA = \$NR_AGENCIA\ PJ * (Veículo_Automotor)$$

Administração Central

Tabelas Centralizadas

- Veículo_Automotor
- Todas as tabelas réplicas e extratos das agências.

6.8 - Detalhes de Implementação

1) Quando a tabela Veículo_Automotor (agências) sofrer alteração em alguns de seus atributos (seguro, licenciamento), estas alterações serão propagadas para a tabela master, de forma síncrona, através do mecanismo de gatilho de banco de dados, vinculado à tabela.

2) As tabelas que são extratos nas agências serão objetos Snapshots definidos em cada agência e atualizados a partir da tabela master. Um serão atualizadas de forma assíncrona, utilizando-se o mecanismo de timestamp. Outras serão atualizadas de forma síncrona, através de gatilho de banco de dados vinculado a cada uma das tabelas.

3) As tabelas que são réplicas nas agências, também serão atualizadas através de gatilho de banco de dados definido nas tabelas centralizadas. As atualizações das réplicas dar-se-ão de forma síncrona.

4) O particionamento da tabela Veículo_Automotor não é suportado pelo SGBD. Este particionamento será realizado via programação.

5) Existem alguns processos do sistema de informação que merecem algumas considerações:

- *Abastecimento do Veículo* : Existem atualmente na empresa 4 agências (Florianópolis, Joinville, Blumenau e Lages) que possuem bomba de abastecimento. O abastecimento de veículos de outras agências nestas bombas deve ser repassado à agência de origem do veículo. Isto é realizado através de um processo distribuído, utilizando-se transparentemente o protocolo 2PC implementado pelo SGBD.

- *Empréstimo de Veículo* : Um veículo de uma agência pode ser emprestado para uma outra agência para suprir necessidades temporárias (época de veraneio, catástrofes regionais, etc). Esta transação é disparada pela gerência de transporte da Administração Central. Na agência de origem, o veículo emprestado recebe a identificação de emprestado e, na agência receptora, o veículo é cadastrado com o status de empréstimo temporário.

No retorno do veículo, os custos no período de empréstimo permanecerão na agência receptora, porém os elementos agregados (peças, equipamentos), os dados necessários ao acompanhamento da vida útil e os índices de uso do veículo também retornarão com o veículo emprestado.

Este processo envolverá 3 sítios : na agência central o processo é disparado, ocorrendo inicialmente a atualização do cadastro centralizado; em seguida é realizada a movimentação do veículo da agência de origem para a agência receptora, ocorrendo em ambas as devidas inserções e modificações nas tabelas Veículo-Automotor, Operação-Veículo e Vinculação-Veículo.

Toda transação se dará de forma distribuída, mantendo-se a integridade do banco de dados distribuído através do uso transparente do protocolo 2PC.

- *Transferência de Veículo* : Em algumas situações um veículo pode ser transferido de forma definitiva de uma agência para outra. O processo é semelhante ao processo de empréstimo. As linhas das tabelas vinculadas ao veículo transferido são copiadas para a agência destino, para efeito de depreciação do veículo, controle de peças, inspeções, etc. Na agência de origem este evento é registrado na tabela Operação-Veículo. A linha na tabela Veículo-Automotor correspondente ao veículo movimentado recebe a identificação de transferido.

- *Cadastro de Funcionários* : O cadastro de funcionários reside no SGBD DB2 no mainframe. Neste banco de dados foi disponibilizada uma tabela log das alterações sofridas neste cadastro com a respectiva data de ocorrência. Inicialmente cada agência fará uma cópia síncrona deste cadastro. As futuras atualizações desta réplica ocorrerão de forma assíncrona, em um período de tempo a ser determinado, de forma automática, consultando apenas a tabela de log no mainframe. A comunicação entre os SGBDs Oracle7 e DB2 dar-se-á através do gateway Transparent Gateway For DB2, da Oracle Corporation.

6.9 - Conclusão

Este capítulo teve como objetivo a definição da forma de distribuição das tabelas do sistema distribuído Controle da Frota de Veículos em cada um dos locais de abrangência deste. A partir dos assuntos de distribuição e dos locais envolvidos, chegou-se ao Plano de Compartilhamento Físico, relacionando-se todas as tabelas e sua forma de distribuição em cada sítio.

Foram apresentados também alguns processos de características distribuídas que o sistema deve implementar de forma transparente, dando a impressão de que se está trabalhando em um grande banco de dados centralizado.

Para armazenar as ocorrências anuais das tabelas e índices, cada sítio deve dispor de 30 MB de espaço em disco. No total o sistema necessitará de 480 MB (16 agências x 30 MB).

Uma vez apresentado o sistema de informação e o ambiente onde foi implementado, ver-se-á no próximo capítulo a análise do modelo implementado com relação ao modelo centralizado e também se verificará a implementação quanto ao atendimento das regras de um banco de dados distribuído ideal, definidas por Christian Date [10].

Capítulo 7

ANÁLISES DO MODELO IMPLEMENTADO

7.1 - Introdução

Duas análises são apresentadas neste capítulo : a primeira enfocará os benefícios obtidos com o modelo distribuído do sistema de informação tendo-se como parâmetro de comparação o mesmo modelo visto de forma centralizada; a segunda análise será desenvolvida verificando-se o quanto o modelo de sistema distribuído implementado atende os 12 requisitos de um banco de dados distribuído ideal [7].

Na análise envolvendo o modelo distribuído e o centralizado, alguns itens foram evidenciados : o tempo de resposta, transmissão de dados, armazenamento, disponibilidade, manutenção, desenvolvimento e documentação, backup e recuperação, funções do DBA e informações gerenciais.

Na segunda análise, verificou-se o modelo distribuído do sistema com relação aos seguintes itens de um banco de dados ideal [7] : autonomia de local, sem dependência de um centro de processamento de dados central, operação contínua, independência de local, independência de fragmentação, independência de replicação, processamento de consulta distribuída, gerenciamento de transação distribuída, independência de máquina, independência de sistema operacional, independência de rede e independência de SGBD.

7.2 - Análise do Modelo Implementado com Relação ao Modelo Centralizado

O Sistema de Informação Controle de Frota de Veículos foi desenvolvido de forma totalmente distribuída, aproveitando-se suas características distribuídas.

Foram selecionados alguns parâmetros considerados importantes para se fazer uma análise satisfatória, mostrando-se os pontos positivos e negativos de uma implementação distribuída de um sistema.

Tempo de Resposta, Transmissão de Dados na Rede, Armazenamento, Disponibilidade, Manutenção do Sistema, Desenvolvimento e Documentação, Backups e Recuperação, Funções do DBA, Informações Gerenciais são os assuntos que serão analisados.

7.2.1 - Tempo de Resposta

Para acessar os dados residentes no computador de grande porte, os usuários das agências regionais e administração central utilizam terminais remotos monocromáticos e linhas de transmissão de 4800 e 9600 bps.

Com a descentralização do processamento de dados para junto de quem realmente os gera e os utiliza, as agências, e com a adoção de banco de dados distribuído, o tempo de resposta tende a melhorar sensivelmente, pois o acesso aos dados se dará localmente (quase que totalmente) através de rede local de 10Mbps.

O acesso local aos dados, aliado à velocidade da rede local, faz com que os aplicativos cliente/servidor do sistema distribuído tenham o tempo de resposta bastante reduzido com relação ao mesmo aplicativo acessando os dados de forma centralizada, através de terminais remotos.

7.2.2 - Transmissão de Dados na Rede

Antes da implementação do sistema Controle de Frota de Veículos descrita neste trabalho, o sistema anterior funcionava praticamente em modo batch e era

bastante limitado em seus benefícios gerenciais. A informação retornava às agências (total de 16) com atraso de meses. Vislumbrar todos os tipos de aplicativos possíveis, todos os tipos de consultas locais a serem realizados, todos os tipos de relatórios gerenciais a serem montados é uma tarefa bastante difícil. Para facilitar esta análise, serão considerados somente os dados envolvidos com os principais aplicativos, espelhando apenas uma parte do mundo real.

Os bits de formatação de telas da família de terminais 3270 serão desconsiderados porque não se tem uma idéia clara de todos os aplicativos e de como seriam seus desenhos de telas.

O tráfego de dados na rede de longa distância com o sistema distribuído ocorrerá nas seguintes situações : atualizações das tabelas acessórias (são praticamente estáticas); dados sumarizados sobre utilização dos veículos em todas as agências (realizados mensalmente); tabelas de empregados (atualizadas a partir de log de tabela no mainframe); transferência e empréstimo de veículo (eventualmente) e abastecimento de veículos de outras agências (eventualmente). Diariamente, em média, estas movimentações têm impacto zero na rede, sendo portanto desconsideradas para análises.

A grosso modo, a quantidade de bytes que estaria trafegando na rede de comunicação de longa distância para a forma centralizada do sistema, analisando-se os principais aplicativos (ver descrição dos aplicativos no capítulo 2) é :

01) Despesas Diversas com Veículos

Consultas : $880 \text{ (bytes)} * 16 \text{ (agências)} = 14.080 \text{ bytes}$

Atualizações : $1.440 * 16 = 23.040 \text{ bytes}$

02) Operações com Veículos

Consultas : $1.170 * 16 = 18.720 \text{ bytes}$

Atualizações : $12.480 * 16 = 199.680 \text{ bytes}$

03) Serviços Realizados pelos Veículos

Consultas : Não existem

Atualizações : $712 * 16 = 11.392 \text{ bytes}$

04) Veículos

Consultas : $95 * 16 = 1.520$ bytes

Atualizações : $23 * 16 = 368$ bytes

Total Diário = 268.700 bytes

Total Mensal = 7 Mbytes

De um modo geral, poder-se-ia dizer que com a adoção do banco de dados distribuído do sistema Controle de Frotas de Veículos, 7 Mbytes de dados deixarão de trafegar mensalmente na rede de longa distância, porque os acessos estarão sendo realizados localmente.

7.2.3 - Armazenamento

O armazenamento particionado da tabela de Veículo-Automotor e suas tabelas agregadas (tabelas de receitas, despesas, acidentes e operações) faz com que não se tenha nenhum acréscimo com relação ao modelo centralizado.

A tabela de Empregados é replicada de forma particionada, não havendo portando acréscimo nos recursos de armazenamento de dados.

As tabelas acessórias, formadas por extratos e réplicas, possuem um volume de dados insignificante com relação às principais tabelas do sistema.

As tabelas contendo dados sumarizados de índices de utilização dos veículos é que realmente ocuparão áreas de armazenamento extras.

$$\begin{aligned} \text{Área Extra Mensal} &= 55 \text{ (tamanho médio da linha) } * \\ & 6684 \text{ (número de linhas) } * \\ & 15 \text{ (número de agências - 1)} \end{aligned}$$

$$\text{Área Extra Mensal} = 5,5 \text{ Mbytes}$$

O total estimado de dados, caso fosse centralizado, é de 78 Mbytes.

7.2.4 - Disponibilidade

O ambiente de grande porte é bastante estável na Celesc em termos de disponibilidade diária dos dados aos usuários finais. Entretanto, uma pane de CPU ou manutenções preventivas e corretivas podem fazer com que todas as agências fiquem impossibilitadas de acessar informações cruciais ao seu dia-a-dia, principalmente dados referentes à arrecadação e faturamento.

Toda manutenção possível de ser planejada pode ser realizada nos finais de semana. A disponibilidade do grande porte poderia ser melhorada caso se duplicasse todos os seus recursos, mas isso teria um custo muito alto.

A disponibilidade dos dados de Controle da Frota de Veículos Distribuído é consideravelmente maior. Os problemas relacionados a falhas, inspeções, correções de linhas e equipamentos são isolados a uma determinada agência não interferindo nas demais. A utilização de Extratos e Réplicas torna esta disponibilidade ainda maior.

7.2.5 - Manutenção de Sistemas

A manutenção dos sistemas distribuídos pode tornar-se um processo difícil. Para manutenções dos bancos de dados distribuídos existem ferramentas de funcionamento centralizado que possibilitam transparentemente o gerenciamento dos bancos distribuídos.

A forma de uso dos dados na agências dar-se-á na modalidade cliente/servidor. Alguns processos de negócios das aplicações residirão no SGBD local na forma de gatilhos de banco de dados e procedimentos armazenados. As manutenções destes processos podem ser realizadas de forma centralizada.

Nas agências os aplicativos ficarão armazenados em um servidor Novell. Os módulos atualizados dos aplicativos, como também novos aplicativos, poderão ser gerenciados também de forma centralizada, utilizando-se a integração entre rede de longa distância e redes locais.

7.2.6 - Desenvolvimento e Documentação de Sistemas

Os aspectos de levantamento de dados, modelagem de dados e projeto lógico global continuam a existir de acordo com a metodologia que a empresa vinha adotando. Serão acrescentados o projeto lógico distribuído e o projeto físico distribuído. Estas novas etapas carecem de alguns estudos e pesquisas para se chegar a uma metodologia que melhor se adapte ao método de desenvolvimento de sistemas já praticado na empresa.

A metodologia de documentação de sistemas sofrerá alterações para inserir as novas etapas, ou seja, projeto lógico e físico de banco de dados distribuído.

7.2.7 - Backup e Recuperação

Quando se implementa Banco de Dados Distribuído o pressuposto básico é que se tenha em cada SGBDD o *log* ativo e também arquivamento automático de *log*.

O gerenciamento de backups é uma das mais importantes funções quando se fala de banco de dados distribuído. Em um SGBD centralizado esta é uma tarefa relativamente simples de ser realizada. Recuperar um SGBD centralizado após falhas graves no sistema de discos, principalmente, é realizado de forma bastante eficiente.

Cada sítio de um ambiente de banco de dados distribuído deve possuir total autonomia em suas funções e gerenciamento. Contudo, quando se fala de banco de dados distribuído, deve-se levar em conta a integridade do banco de dados global. Em caso de falhas graves em um sítio, o restabelecimento do banco de dados neste sítio deve fazer com que o banco de dados distribuído retorne íntegro. Para tanto, deverá ocorrer uma comunicação entre os sítios (de forma centralizada ou não) para negociar o que fazer para manter o sincronismo entre os bancos.

O gerenciamento dos backups do ambiente distribuído será realizado por uma ferramenta contratada que irá contornar os problemas citados acima.

7.2.8 - Funções do DBA

Com o advento dos Sistemas Gerenciadores de Bancos de Dados Distribuídos, novas funções surgiram para o Administrador de Banco de Dados, já descritas no final do capítulo 1.

Se a função de administrar o ambiente distribuído for centralizada, dependendo do número de sítios, esta tarefa pode se tornar bastante laboriosa. O auxílio de ferramenta de administração centralizada de SGBDD pode ser a grande solução.

Pela teoria, em um banco de dados distribuído ideal, cada sítio deve possuir total autonomia, inclusive no aspecto de gerenciamento administrativo. Contudo, na prática, existem vários inconvenientes que poderão fazer com que as funções de um DBA seja realizadas de forma centralizada.

7.2.9 - Informações Gerenciais

As informações gerenciais freqüentemente utilizadas pelo corpo gerencial da empresa nas tomadas de decisões e acompanhamento de desempenho da empresa, no modelo centralizado, usando o SGBD DB2, são muito fáceis de serem obtidas. O processamento se dá em um único local e em um único SGBD.

Com o banco de dados distribuído, obter informações gerenciais dará um pouco mais de trabalho. As informações corporativas estão no computador de grande porte no SGBD DB2. As informações distribuídas estão armazenadas em máquinas RISC/6000 em SGBD Oracle7. Uma informação gerencial pode gerar consultas em 16 sítios diferentes em todo o estado e também gerar acesso ao computador de grande porte, para então processar e disponibilizar o resultado.

7.3 - Análise do Modelo Implementado quanto ao Banco de dados Ideal

Para realizar esta análise escolheu-se os 12 requisitos de Banco de Dados Distribuído Ideal estabelecido por Christian J. Date. Estas regras estão descritas no início do capítulo 1 [7].

7.3.1 - Autonomia Local

Esta regra é plenamente atendida pelo modelo implementado, não existindo portanto um dicionário, controlador de atividades ou detector de bloqueio mútuo centralizado.

A dependência de um nó central nos casos de extratos e réplicas é justificada pela melhoria de performance, principalmente.

As funções de DBA ocorrem de forma centralizada por motivos econômicos e técnicos.

7.3.2 - Sem Dependência de um Centro de Proc. de Dados Central

Funções tais como gerenciamento de dicionário de dados, processamento de pesquisa, controle de concorrência e controle de recuperação devem ser tratadas de forma distribuída. Uma pane no sítio centralizador dessas funções poderia acarretar um colapso no sistema como um todo.

O SGBD Oracle7 não implementa catálogo de dados distribuído e nem controle de concorrência distribuído. O controle de concorrência ocorre através do mecanismo de timeout. A recuperação de transações com falhas ocorre de forma transparente através do protocolo de duas fases implementado pelo SGBD.

O SGBD Oracle7 possibilita a replicação do catálogo, controlada de forma automática.

7.3.3 - Operação Contínua

A adição de um novo sítio ou a manutenção de um determinado sítio não deve tornar o sistema inoperante.

No modelo implementado esta regra é plenamente atendida. A todo instante todos os acessos aos dados estarão sendo realizados localmente.

7.3.4 - Independência Local

Esta regra não é atendida pelo fato de não se ter catálogo de dados distribuído, onde ter-se-ia a definição do esquema de fragmentação dos bancos de dados distribuídos.

A transparência de localização é conseguida através do uso de visões, sinônimos e database link. A aplicação/usuário deve trabalhar com um destes conceitos para manipular transparentemente os dados. O database link é uma definição a nível de catálogo do sistema onde se registra o caminho de acesso ao dado. Nas consultas ou atualizações distribuídas de tabelas basta informar o database link da mesma.

A mudança de local de réplicas e fragmentos não será sentida pelo usuário/aplicação, porque basta o DBA acertar o novo caminho no database link.

7.3.5 - Independência de Fragmentação

A fragmentação de dados não é suportada pelo SGBD Oracle7. Tanto a definição quanto a manipulação de dados fragmentados só são possíveis utilizando-se programação e comandos SQL.

A localização de cada fragmento é feita via database link.

A fragmentação horizontal, vertical e mista pode ser implementada manualmente. Consultas e atualizações de fragmentos são realizadas indicando-se a localização do(s) fragmento(s).

7.3.6 - Independência de Replicação

Esta é uma área onde os SGBDs comerciais estão evoluindo bastante e onde está ocorrendo bastante investimentos.

O Oracle7 implementa réplica, denominada de Snapshot, somente de leitura ou atualizáveis, inclusive formando clusters para manter relacionamentos. Estas réplicas possuem mecanismos automáticos de atualização síncrona e assíncrona. As réplicas podem ser completas ou parciais. Quando a forma é parcial, deve haver a definição de um log para a tabela master.

Como não existe a figura de catálogo distribuído, a localização das réplicas no sistema não é possível de ser feita de forma automática.

Não é possível ao otimizador de consulta fazer uso de réplicas pelo simples fato de localmente não saber onde se encontram todas as réplicas de uma determinada tabela.

7.3.7 - Processamento de Consulta Distribuída

Uma das razões pela qual o SGBD Oracle7 não atende este requisito é o fato de não implementar esquema de fragmentação e catálogo distribuído. Como estabelecer estratégias de acesso utilizando-se réplicas se não existem informações globais de definição e localização?

Para que o otimizador de consulta possa utilizar dados estatísticos sobre as tabelas é necessário periodicamente executar um utilitário denominado *Analyze*, para que as informações a respeito de número de linhas, tamanho médio de colunas chaves, índices, etc, sejam registradas no catálogo.

Em uma consulta distribuída envolvendo junção, o sítio onde é inicialmente processada a transação pode consultar previamente os sítios envolvidos na consulta para fazer a menor movimentação de dados possível e, nos sítios locais, explorar o otimizador de consulta local.

7.3.8 - Gerenciamento de Transações Distribuídas

Dois aspectos importantes do gerenciamento de transações distribuídas são o controle de concorrência e o controle de recuperação. O controle de

recuperação é plenamente atendido pelo Oracle7 através de um transparente mecanismo de confirmação de duas fases. Com este SGBD, uma transação distribuída é totalmente confirmada ou totalmente desfeita.

O Oracle7 não implementa controle de concorrência global. O controle de concorrência utilizado é o implemento pelos SGBDs locais, ou seja, o de bloqueio. A detecção de bloqueio mútuo global é resolvida pelo mecanismo de timeout.

7.3.9 - Independência de Máquina

Com esta regra o que se espera é que o SGBDD rode nos principais hardwares de mercado e em várias plataformas, desde computadores pessoais até computadores de grande porte. Este item também é atendido.

7.3.10 - Independência do Sistema Operacional

O sistema operacional deve ser transparente ao SGBDD. Novamente, leia-se os sistemas operacionais de mercado.

O Oracle7 atende também esta regra para os principais sistemas operacionais de mercado, podendo rodar sob DOS/Windows, UNIX, OS/2 , MVS/ESA, etc.

7.3.11 - Independência de Redes de Comunicação

Os SGBDDs devem se comunicar utilizando-se de todos os meios de comunicação de dados e protocolos existentes. O meio de comunicação deve ser um item transparente ao banco de dados distribuído. Esta regra também está de acordo em nosso modelo implementado.

7.3.12 - Independência de SGBD

Deve-se verificar neste requisito as possibilidades de transação distribuída, consulta distribuída, definições de réplicas, fragmentos, etc, entre SGBDDs homogêneos e heterogêneos.

O Oracle7 implementa satisfatoriamente esta regra, possuindo acesso bidirecional com alguns SGBDs a nível de leitura e atualização distribuída.

7.4 - Conclusão

Neste capítulo analisou-se alguns itens referentes ao desenvolvimento, uso e manutenção de um sistema de informação e os aspectos de centralização e distribuição de dados. Nesta análise não se considerou o ambiente ideal para a implementação do modelo distribuído, mas o ambiente atual da Celesc. Para tornar o ambiente mais propício ao uso de banco de dados distribuído na empresa, ter-se-ia que reestruturar totalmente o ambiente de comunicação de dados, refazer a metodologia de desenvolvimento de sistemas, melhorar a capacidade de processamento dos computadores RISC/6000, atualizar as versões do sistema operacional (AIX) e do SGBD (Oracle7) e, em alguns aspectos, reformular o modelo gerencial adotado pela empresa.

Na análise do modelo implementado quanto a um banco de dados distribuído ideal [7], verificou-se o SGBDD Oracle7 e as opções de banco de dados distribuído que ele implementa. Como pontos fortes do SGBDD destacam-se a replicação de dados e gerenciamento de transações distribuídas. Como pontos negativos do SGBDD Oracle7 tem-se a não implementação de controle de concorrência global (o controle de concorrência não ocorre de forma global), esquema de fragmentação e otimização adequada de consultas distribuídas.

Para as necessidades distribuídas do sistema de informação Controle da Frota de Veículos, o SGBDD Oracle7 possibilita quase que totalmente as implementações distribuídas.

Capítulo 8

CONCLUSÃO

Neste trabalho foi apresentado o estudo realizado sobre Bancos de Dados Distribuídos. Descreveu-se a implementação do sistema de informação distribuído Controle da Frota de Veículos da Celesc e os resultados das análises realizadas confrontando-se o banco de dados distribuído desenvolvido com os requisitos funcionais básicos de um banco de dados distribuído ideal. Analizou-se também alguns itens comuns a um banco de dados distribuído e centralizado, apresentando-se os pontos fortes de uma implementação distribuída.

Na descrição do sistema de informação distribuído, a nível de projeto lógico e físico, utilizou-se uma metodologia desenvolvida pela empresa Forma Informática. Nesta metodologia, os passos para se chegar ao modelo lógico e físico do sistema consistem no preenchimento de matrizes bidimensionais. No projeto lógico, a matriz resultante relaciona a forma de armazenamento de cada entidade com cada assunto de informação do sistema. O projeto físico detalha para cada entidade do sistema a forma de armazenamento em cada um dos sítios do banco de dados distribuído.

Na implementação do sistema de informação distribuído utilizou-se os recursos físicos da Celesc. O SGBDD Oracle⁷ é o responsável pela gerência do banco de dados distribuído do sistema implementado, cabendo ao SGBD DB2 a gerência dos dados corporativos. Os recursos utilizados foram os mais diversos: equipamentos RISC/6000, micro PENTIUM, rede WAN, rede local, além de outros.

Este trabalho apresenta algumas limitações em virtude do estágio atual das implementações distribuídas dos SGBDD's comerciais, da falta de uma metodologia clara e abrangente de desenvolvimento de sistemas de informação

distribuídos, das limitações dos recursos de comunicação de dados da empresa e também pelas indefinições na descentralização dos processos administrativos.

O fato do SGBDD Oracle7 não implementar fragmentação de dados implica em programação via gatilhos e procedimentos remotos para que esta característica possa ser utilizada na definição do banco de dados distribuído. Mesmo assim, dependendo do tipo de fragmentação, a complexidade pode ser tão grande que inviabilize sua implementação, tendo-se que optar então por replicação de dados.

Por não possuir um catálogo de dados distribuído, o SGBDD Oracle7 não tem como identificar de forma automática a localização das entidades de dados remotas e nem como determinar o melhor caminho de acesso para se chegar aos fragmentos e réplicas. Em cada catálogo local deve existir a definição de cada entidade remota e o caminho para se chegar até ela. Com isto, a Otimização de uma consulta distribuída fica bastante prejudicada.

Um item que pesou bastante na implementação foi a utilização de linhas de comunicação de baixa velocidade. Com o seu uso, procura-se ao máximo evitar o tráfego de dados em horários de pico e, com isto, o uso desnecessário de réplicas torna-se uma prática bastante utilizada, gerando assim complexidades que o sistema não possui, além de se ter gastos desnecessários de armazenamento.

Um outro fator que prejudicou bastante a implementação distribuída do sistema foi a forma de descentralização gerencial atualmente vigente em algumas áreas da Celesc, onde a dependência gerencial de uma agência regional com a administração central ainda é bastante forte, apesar de ser tecnicamente desnecessária, valendo-se tão somente da cultura existente. Esta anomalia gera novas funções ao sistema, de cunho burocrático, que em nada contribui para a eficiência no gerenciamento da informação.

Estes foram alguns pontos que tornaram o trabalho um pouco mais limitado em sua funcionalidade, e pobre em alguns aspectos de implementação.

Com o conhecimento que se tem dos sistemas de informação implantados na Celesc, com o estudo realizado sobre Bancos de Dados Distribuídos e com a experiência de uma implementação de sistema distribuído, pode-se afirmar que as características distribuídas apresentadas pelo SGBDD Oracle7 atendem praticamente todas as necessidades de distribuição de dados da empresa. Este é um ponto bastante forte, cuja conclusão foi proporcionada pelo desenvolvimento do presente trabalho.

O estudo da metodologia de desenvolvimento de sistemas de informação distribuídos utilizado na implementação também foi bastante relevante no sentido de se ter, a partir daí, um conhecimento razoável para se criar na Celesc uma metodologia que comporte esta nova forma de desenvolvimento de sistemas. A

confrontação entre o banco de dados distribuído implementado e os requisitos funcionais básicos ideais proporcionou um quadro das reais implementações distribuídas do SGBDD Oracle7 e, com isto, propiciou conclusões a serem consideradas no desenvolvimento de novos sistemas na empresa.

A comparação entre a implementação distribuída com relação ao modo centralizado serviu para demonstrar uma série de vantagens do uso de sistemas distribuídos na empresa.

A estrutura de comunicação de dados na Celesc, que hoje é bastante precária, está passando por uma reformulação em termos de capacidade de transmissão, roteamento e integração das agências regionais. A descentralização dos processos administrativos é algo que irá aos poucos se consolidando.

A metodologia de desenvolvimento de sistemas distribuído é, com certeza, o grande foco de estudos para que a prática de desenvolvimento de sistemas utilizando banco de dados distribuído torne-se uma realidade na empresa.

Novos trabalhos terão que ser realizados na empresa, vinculando-se o desenho, projeto e implementação de sistemas de informação distribuídos orientados a objetos, pois a Oracle Corporation está anunciando um novo SGBDD orientado a objetos para o início do segundo semestre de 97, e a Celesc tem interesse nesta tecnologia.

A nível teórico, existem alguns assuntos interessantes para acompanhamento a nível acadêmico. São eles: otimizador de consultas distribuídas; resolução de deadlocks globais; implementação do esquema de fragmentação; algoritmos de controle de concorrência em ambiente distribuído; protocolos de resolução de transações distribuídas e dicionário de dados distribuído.

ANEXOS

Anexo1 - Relação das entidades do sistema de informação Controle da Frota de Veículos

ACID-VEICULO-DANO

Relaciona os danos do veículo causado por acidente.

ACIDENTE-VEICULO

Acidente envolvendo veículo da empresa.

ATIVIDADE-USO-VEIC

Descrição dos motivos da utilização do veículo.

BOMBA-COMB-CELESC

Bomba de abastecimento de combustível localizada nas dependências da Celesc.

CALENDARIO

Identifica os dias, semanas e meses do ano civil.

CARAC-VEICULO

Especificação do tipo, modelo e complemento de veículo automotor.

CARGO

Estrutura de cargos da Celesc.

CAUSA-ACID-VEIC

Razão, motivo ou origem do acidente com veículo.

CLND-FREQ-TRABALHO

Calendário da empresa contendo feriados municipais, compensação, ponto facultativo e liberalidade da Empresa.

CONDUTOR-VEICULO

Funcionário da Celesc designado para conduzir veículo, no transporte de materiais, pessoas ou serviços.

CONVERSAO-CMB-VEIC

Adaptação no motor, para substituir o tipo de combustível utilizado pelo veículo.

COTACAO-MOEDA

Contém cotações de moedas, por exemplo, dólar.

CST-HOMEM-HR-TRANP

Custo do homem-hora da função transporte.

DESP-MANU-VEIC-TOT

Gasto total de uma Ordem de Serviço para Veículo (OSV) com serviços de reparo, manutenção, revisão, adaptação e outros, realizados em um veículo. É composto pela mão-de-obra própria e serviços de terceiros.

DESP-PESSOAL-TRANP

Gastos com salários, remunerações, encargos sociais e benefícios diversos relativos aos empregados relacionados direta ou indiretamente com a função transporte.

DESPESA-DVRS-VEIC

Gastos diversos (sem OSV) com veículos da empresa, como abastecimento de combustível, realizado em bomba própria ou de terceiros, licenciamento, seguro, etc.

DESPESA-INDR-TRANP

Outros gastos da função transporte (exceto despesa com pessoal) não individualizando o veículo.

DESPESA-MANUT-VEIC

Gastos através de uma OSV com serviços de reparo, manutenção, revisão, adaptação e outros, realizados em um veículo. Inclui peças, mão-de-obra própria e serviços de terceiros.

DESPESA-TRANSPORTE

Todas as despesas diretas ou indiretas da função transporte.

DSP-DUPLA-FUNCAO

Gastos com pagamento de dupla função ao empregado autorizado a conduzir veículo.

EQPTO-ACOPL-VEIC

Equipamentos e acessórios instalados no veículo, tais como: guindaste, escada giratória, ferramentas, etc.

EQPTO-VEICULO

Aparelhos, ferramentas e acessórios utilizados para adaptar o veículo, tornando-o apto a realizar tarefas pertinentes a Celesc.

ESPDD-FORN-REGL

Descreve a especialidade do fornecedor de peças e serviços para veículos.

EXTR-TRNP-ATVD-RGL

Resumo de todas as ocorrências de operação, serviço, despesa, custo e outros indicadores de transporte, totalizado por tipo de atividade de utilização na regional.

EXTR-TRNP-CRT-VEIC

Resumo de todas as ocorrências de operação, serviço, despesa custo e outros indicadores de transporte, totalizado por característica de veículo.

EXTR-TRNP-EMPRESA

Resumo de todas as ocorrências de operação, serviço, despesa, custo e outros indicadores de transporte utilizados na Celesc.

EXTR-TRNP-REGIONAL

Resumo de todas as ocorrências de operação, serviço, despesa, custo e outros indicadores de transporte utilizados na Celesc, totalizado por agência.

EXTR-TRNP-USU-VEIC

Resumo de todas as ocorrências de operação, serviço, despesa, custo e outros indicadores de transporte utilizados na Celesc, totalizado por órgão usuário de veículo.

EXTR-TRNP-VEICULO

Resumo de todas as ocorrências de operação, serviço, despesa, custo e outros indicadores de transporte utilizados na Celesc, totalizado por veículo.

EXTR-TRNP-VINC-VEI

Resumo de todas as ocorrências de operação, serviço, despesa, custo e outros indicadores de transporte utilizado na Celesc, totalizado por órgão de vinculação do veículo.

FORNECEDOR-REGL

Fornecedor de materiais, serviços, energia, aluguéis, etc.

MARCA-VEICULO

Marca de veículo segundo seu fabricante. Geralmente é o próprio nome do fabricante (FORD, FIAT, GM, etc).

MOEDA

Contém os nomes das moedas utilizadas pela Celesc.

MUNICIPIO

Tabela dos municípios catarinenses.

NATZ-DANO-ACD-VEIC

Condição , estado, prejuízo ou danificação do veículo e seus equipamentos, resultante de acidente.

OPER-VEIC-TERC

Informações sobre a utilização de veículos de terceiros, por contrato de locação, veículo de empregado e táxi, no deslocamento de pessoas e/ou materiais.

OPERACAO-VEICULO

Informações sobre a utilização de veículo da frota da Celesc no deslocamento de pessoas e/ou materiais.

ORGAO-ADM-CELESC

Representa a estrutura administrativa da Celesc, indicando os órgãos da administração Central e Regional.

PRECO-COMB-VEIC

Preço dos combustíveis no último dia do mês. Utilizado no cálculo dos custos da função transporte.

RECEITA-INDR-TRANP

Descreve receitas gerais da função transporte como um todo, sem associá-las a um veículo em especial.

RECEITA-TRANSPORTE

Todas as receitas diretas ou indiretas da função transporte.

RECEITA-VEICULO

Devolução ou ressarcimento feito por terceiros ou por empregados, em função de despesas com manutenção em veículos que sofreram acidentes, multas ou outras despesas pagas pela Celesc.

SERVICO-VEICULO

Descreve todo reparo, ajuste, adaptação, etc, realizado em um veículo para deixá-lo em condições de uso (manutenção, instalação de equipamentos, lubrificações, revisões, etc).

VEICULO-AUTOMOTOR

Equipamento automotor de transporte e serviço utilizado na Celesc (automóvel, caminhão, caminhonete, motocicleta, etc).

VINCULACAO-VEICULO

Cessão de veículo para um determinado órgão da Celesc.

VITIMA-ACID-VEIC

Pessoa, empregada ou não, lesada em acidente com veículo da Celesc.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] - **ADMINISTRATION oracle7 with distributed option.**
Belmont,USA : Oracle Corporation, 1995.
- [2] - **DB2 family solutions.** San José : IBM Corporation, 1994
- [3] - **DB2 for mvs/esa.** San José,USA : IBM Corporation, 1994.
- [4] - BELL, David ; GRIMSOM, Jane. **Distributed database systems.**
Workingham,England : Addison-Wesley, 1992. 410 p.
- [5] - CASANOVA, Marco Antônio. **Princípios de sistemas de gerência de banco de dados distribuídos.** Rio de Janeiro : Campus, 1985. 355 p.
- [6] - CERI, Stefano ; PELAGATTI, Giusepper. **Distributed databases : principles & systems.** New York : McGraw-Hill, 1992. 410 p.
- [7] - CERICOLA, Osvaldo Vicente. **Banco de dados relacionais e distribuído.**
Rio de Janeiro : Livros Técnicos e Científicos, 1991. p 213-243.
- [8] - COMITÊ de gestão empresarial. (COGE). **Processamento distribuído em multiplataforma.** Rio de Janeiro : [sn], 1995.
- [9] - **CONGRESSO nacional de novas tecnologias e aplicações em banco de dados. 7. : 1996 : São Paulo.**

- [10] - DATE, C. J. **Introdução a sistemas de banco de dados**. Rio de Janeiro : Campos, 1989. 513 p.
- [11] - **DISTRIBUTED relational database architecture**. San José,USA : IBM Corporation, 1995. 477 p.
- [12]- HOPPER, Teresa (Editor). **Distributed relational database architecture : connectivity guide**. San José ,USA : IBM Corporation, 1995. 477 p.
- [13] - MARTIN, James. **Strategic data planning methodologies**. Savant Research Studies,USA, 1980. 275 p.
- [14] - **ORACLE7 for mvs system administration guide**. Belmont,USA : Oracle Corporation, 1996.
- [15] - **ORACLE7 server administrator's guide**. Belmont,USA : Oracle Corporation, 1996.
- [16] - **ORACLE7 server concepts**. Belmonte,USA : Oracle Corporation, 1996.
- [17] - **ORACLE7 server distributed systems**. Belmonte,USA : Oracle Corporation, 1996.
- [18] - **ORACLE7 server migration**. Belmont,USA : Oracle Corporation, 1996.
- [19] - **ORACLE7 transparent gateway for db2**. Belmont,USA: Oracle Corporation, 1996.