

UNIVERSIDADE FEDERAL DE PERNAMBUCO

**GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO
CENTRO DE INFORMÁTICA**

2004.2

**MODELAGEM INCREMENTAL DE
UM AMBIENTE DE DATA WAREHOUSING:
UMA ABORDAGEM DA ARQUITETURA BOTTOM-UP**

**PROPOSTA DE TRABALHO DE
GRADUAÇÃO EM
BANCO DE DADOS**

**Aluno: José de Anchieta Couto Caraciolo Netto, jaccn@cin.ufpe.br.
Orientador: Fernando da Fonseca de Souza, fdfd@cin.ufpe.br.**

Recife, 8 de Novembro de 2004.

Índice

1. CONTEXTO.....	3
2. PROJETO.....	4
3. CRONOGRAMA.....	5
4. REFERÊNCIAS	6
5. ASSINATURAS.....	7

1. Contexto

As inovações mundiais na tecnologia de hardware e software, combinadas com as conseqüentes mudanças das visões de negócios, levaram as corporações a se preocuparem com o controle de suas informações. Os Sistemas de Suporte à Decisão que compõem o Ambiente de Apoio a Decisão, tornaram-se cada vez mais solicitados, passando a ser peças chaves para análises estratégicas. Porém, a propagação de sistemas nas empresas causou um aumento no volume das informações processadas e uma maior distribuição destas informações, acarretando uma complexidade em seu gerenciamento para o Ambiente de Apoio a Decisão. A complexidade do gerenciamento de informações resultou em uma necessidade de mudanças, atendida com a tecnologia de Data Warehousing. Esta tecnologia está representada na figura do Data Warehouse, uma base de dados capaz de manipular um grande volume de informações com bom desempenho. A abordagem de Data Warehouse permitiu aperfeiçoar a gerência, o controle e o acesso aos dados.

Apesar da crescente absorção da tecnologia e do considerável número de referências relacionadas ao assunto, a tecnologia de Data Warehousing pode ser considerada muito recente. Normalmente, são apresentadas técnicas e formas de transformações empregando a modelagem "Top-down", considerada a modelagem padrão, que primeiramente desenvolve o modelo do DW para depois modelar os Data Marts [INM1997]. A própria fonte bibliográfica utilizada como motivação para realizar este trabalho [CIF2002] considera, como uma forma adicional de arquitetura, a possibilidade do sistema WebD²W utilizar Data Marts como unidade de distribuição, utilizando para isto a abordagem Top-down, uma vez que o processo de distribuição deste sistema é realizado a partir do Data Warehouse global e origina diversas unidades de distribuição.

Entretanto, essa modelagem representa um processo custoso e pouco prático, não apresentando, normalmente, bons resultados, devido ao tempo gasto na sua implementação. As empresas, em sua grande maioria, vêm adotando o desenvolvimento de Ambiente de Data Warehousing empregando uma abordagem incremental, também conhecida como "Bottom-up" [KIM1998]. Essa abordagem permite justificar o custo do desenvolvimento deste ambiente pela rápida apresentação de resultados, atraindo a atenção e investimento por parte dos usuários finais.

2. Projeto

O objetivo deste trabalho de graduação é estabelecer um conjunto de técnicas que permita realizar a modelagem de dados de um Ambiente de Data Warehousing de forma incremental, a partir dos modelos de dados existentes no ambiente operativo. A modelagem do ambiente será realizada a partir da criação de Data Marts, por um processo de derivação dos modelos de dados do ambiente operativo, e sua posterior integração ao Data Warehouse, caracterizando assim uma arquitetura Bottom-up.

Para tornar a seqüência das técnicas mais clara demonstraremos todas as etapas da metodologia de desenvolvimento incremental [SEL2001] [MCG1998]. Esta metodologia se baseia na divisão da modelagem do ambiente em quatro fases:

- 1- Análise dos modelos de dados existentes no ambiente operativo.
- 2- Transformação do modelo de dados, até então orientado a processo, em um modelo de dados mais próximo ao negócio.
- 3- Derivação de um ou mais modelos dimensionais para o Data Mart, a partir do modelo de dados resultante da fase anterior.
- 4- Integração do modelo de dados do Data Mart ao modelo de dados do Data Warehouse, finalizando desta forma a modelagem do Ambiente de Data Warehouse.

Para a aplicação das técnicas propostas neste trabalho será desenvolvido detalhadamente um ambiente de Data Warehousing, determinando-se dois Data Marts a serem construídos, bem como sua prioridade de desenvolvimento e a integração de cada um ao Data Warehouse. Após a criação do ambiente, será desenvolvido um aplicativo simples para exibição dos dados retornados dos parâmetros de consultas definidos pelo usuário, através da geração de relatórios. A linguagem utilizada para implementação será Java, em conjunto com a tecnologia Jasper Reports para a geração dos relatórios. Seguindo a filosofia *open source*, o banco a ser utilizado será o PostgreSQL. A IDE escolhida para o desenvolvimento é o Eclipse e para a geração dos *templates* dos relatórios, o iReport - ambas software livre.

3. Cronograma

O cronograma abaixo demonstra algumas datas para as atividades chaves do processo de desenvolvimento do trabalho de graduação:

Atividade	Mês														
	Novembro			Dezembro			Janeiro			Fevereiro			Março		
1. Estudo detalhado das técnicas de modelagem dos dados.															
2. Modelagem de um caso real de dois Data Marts e respectivas integrações, utilizando as técnicas selecionadas a partir do estudo realizado.															
3. Implementação da modelagem realizada e testes.															
4. Elaborar o relatório.															
5. Elaborar a apresentação.															
6. Defesa do Trabalho de Graduação.															

4. Referências

- [CIF2002] CIFERRI, C. D. A., Distribuição dos Dados em Ambientes de Data Warehousing: O sistema WebD²W e Algoritmos voltados à fragmentação horizontal dos dados. Tese de Doutorado, Universidade Federal de Pernambuco, Recife, PE, BR, 2002.
- [SEL2001] SEEL, D., Uma Arquitetura para Distribuição dos Componentes Tecnológicos de Sistemas de Informações Baseados em Data Warehousing. Dissertação de mestrado, Universidade Federal de Santa Catarina, Florianópolis, SC, BR, 2001.
- [KIM1998] KIMBALL, Ralph, REEVES, Laura, ROSS, Margy, THORNTHWAITE, Warren. The Data Warehouse Lifecycle Toolkit - Expert Methods for Designing, Developing and Deploying Data Warehouses. New York: John Wiley & Sons, Inc., 1998.
- [MCG1998] MCGUFF, Frank. Designing the Perfect Data Warehouse - Data Modeling for Data Warehouse. Disponível na INTERNET via <http://member.aol.com/fmcguff/dwmodel/perfect.htm>. Arquivo consultado em 1998.
- [INM1997] INMON, W. H., Como Construir o Data Warehouse. Campus, Rio de Janeiro, 1997.

5. Assinaturas

Fernando da Fonseca de Souza
Orientador

José de Anchieta Couto Caraciolo Netto
Aluno