



Modalidade do trabalho: Relato de experiência
Evento: 2011 SIT - I Seminário de Inovação e Tecnologia

DATA WAREHOUSE APLICADO A CORRETORA DE SEGUROS¹

Eduardo Ferreira Da Silva², Regis Schuch³, Alencar Machado⁴.

¹ Resumo resultante do Trabalho de Conclusão de Curso da Unijui.

² Estudante do Curso de Ciência da Computação do Departamento DCEENG. eduardo.ferreira@unijui.edu.br.

³ Bacharel em Sistemas de Informações, Unijui. regis.schuch@gmail.com.

⁴ Professor do Departamento de DCEENG Unijui. alencar.machado@unijui.edu.br.

Resumo

A grande quantidade de informações geradas pela tecnologia permite o armazenamento dos mais diversos tipos de dados, com isso acarretando em bases cada vez maiores. O presente trabalho apresenta o uso da tecnologia de modelagem dimensional de um *Data Warehouse*, esta tecnologia foi utilizada sobre uma base de dados oriundos da Corretora de Seguros Karla, com a finalidade de produzir informações gerenciais e de apoio a decisão. Para o desenvolvimento do trabalho foram necessários softwares compatíveis com as tecnologias, logo, foi selecionada a ferramenta *Analysis Services Visual Studio* que está inserida no pacote de funcionalidades do *Microsoft SQL Server 2005* para a modelagem de *Data Warehouse*, e para a obtenção das bases de dados relacionais foi utilizado o IBM DB2.

Palavras-chave: Data Warehouse, Seguros, Sistemas de Apoio a Decisão.

Introdução

A evolução da tecnologia possibilitou que empresas de médio e grande porte possam empregar de forma integral a utilização de sistemas informatizados, para a realização de seus processos mais importantes, com isso, informações registradas ao longo do tempo resultam em um grande volume de dados relacionais. [1] Nós estamos nos afogando em informações, e sedentos de conhecimento. Porém, na prática a utilização dos dados armazenados acaba restringindo-se apenas para simples consultas de informações pertencentes a clientes e negócios, enquanto há uma possibilidade de trabalhar os dados com a finalidade de produzir informações gerencias.

Consequentemente, quanto maior a base de dados, maior a complexidade de análise desses dados, por esta razão o uso de métodos ou técnicas manuais não possibilita capturar informações úteis para gestão de negócios em longo prazo, diante destas necessidades, sejam em ambientes comerciais ou públicos, a técnica de Data Warehouse (DW), tem como propósito organizar os dados e gerar informações não triviais e escondidas a partir de grandes bases de dados. O banco de dados de um DW deve ser projetado para processamento analítico on-line, onde são priorizados a desempenho na recuperação das informações, orientado a





Modalidade do trabalho: Relato de experiência
Evento: 2011 SIT - I Seminário de Inovação e Tecnologia

análise e processos de decisões para usuários finais através de uso de ferramentas especialmente desenvolvidas para o cruzamento multidimensional dos dados [5].

O processo de DW, que em português significa armazém de dados, busca trazer a ideia de centralização e organização das informações em função do tempo e orientada a assuntos, de modo geral o principal objetivo de um DW é proporcionar ao gerente uma flexibilidade de gerar diversos tipos de relatórios e gráficos definidos por seus critérios de refinamento, esses baseados em fatos históricos, e não em intuições ou especulações, por esta razão, torna-se possível uma análise exploratória mais eficaz das informações, portanto essa informação trabalhada pode atuar de forma decisiva como diferencial no processo de tomada de decisões [4].

O presente resumo tem por objetivo principal estudar a aplicação do processo de DW sobre dados de corretora de seguros, com a finalidade de manter os dados sumarizados, periódicos e descritivos. Apresentamos uma revisão sobre *data warehouse*, conceitos sobre seguros, seus tipos de dados e informações que são trabalhadas no seu ramo de negócio, o desenvolvimento do DW e apresentação dos resultados, para finalizar apresentamos as conclusões do trabalho.

Metodologia

Data Warehouse

[3] Define o conceito como um conjunto de dados baseado em assuntos, integrado, não volátil e variável em relação ao tempo, de apoio às decisões gerenciais. [4] DW é uma copia dos dados especialmente estruturados para facilitar o processo de análise, consulta e geração de relatórios.

Portando, partindo dos conceitos fundamentais dos estudiosos, podemos definir como um armazém de dados históricos organizados, limpos e sumarizados, com o objetivo de auxiliar no processo de decisão em ambientes de gestão. O DW pode auxiliar o processo de tomada de decisões, porém, o sucesso dependerá da interatividade e qualificação pessoal do grupo que utilizará a tecnologia.

Para o desenvolvimento de um DW é necessário entender suas fases, estas denominadas: Seleção, Processo de ETC (Extração, Transformação e Carga) e Modelagem Multidimensional. Nas próximas subsecções explicaremos as fases para desenvolvimento de um DW [4].

Seleção

Para o sucesso de um DW é necessário à coleta dos dados, estes geralmente oriundos de mais de um ambiente operacional, levando em consideração que é muito comum encontrar diversas fontes de dados, portanto, no processo de desenvolvimento do DW é necessário que seja desenvolvido um modelo único de dados com a finalidade de agrupá-los [3].

Processo de ETC

O processo de ETC (Extração, Transformação e Carga), é o mais trabalhoso e complexo, demanda a maior parte de tempo do projeto, envolve tarefas como a movimentação dos dados brutos para o DW, limpeza, conversão e carga [2]. No momento em que os dados



Modalidade do trabalho: Relato de experiência
Evento: 2011 SIT - I Seminário de Inovação e Tecnologia

passam a integrar o DW, eles devem passar por um processo de integração em que é preciso usar regras de conversão, ordenação e agregação, unificar entidades semelhantes, padronizar tamanhos, termos, valores etc. [4] O processo de ETC pode envolver os seguintes passos:

- Limpeza dos dados: É comum que na obtenção de dados de bases antigas e desconhecidas, encontrarmos inconsistências. É nessa fase onde os erros de atributos e inconsistências dos dados de diversas tabelas são realizadas;
- Conversão de tipo de dados: Os dados oriundos de fontes diversas podem ter diferentes tipos, estes devem ser convertidos em um padrão único antes de ser integrados ao DW;
- Agregação: Sumarização dos dados observando o nível de granularidade.

Modelo Multidimensional

Os dados contidos em um DW estão organizados em um modelo multidimensional. [3] O modelo multidimensional corresponde a uma maneira intuitiva de organizar dados, facilmente entendida pelo usuário, que consiste numa tabela final denominada tabela fato, dimensões ligadas a ela e um conjunto de métricas descritas por aspectos do negócio.

Granularidade

A granularidade é um dos principais aspectos em um DW, sendo que ela vai definir o nível de detalhamento dos dados que está disponível para o usuário, a decisão sobre qual granularidade será aplicado ao projeto é de suma importância, uma vez que, a escolha de um nível de granularidade inadequada pode comprometer e até inviabilizar o uso do DW [5].

Resultados e Discussão

Seguros e Estudo de Caso

Antes de começar a desenvolver o DW, vamos entender o ambiente onde ele será realizado, e alguns conceitos básicos que são fundamentais na área, sendo que o primeiro conceito será de seguro, esse que pode ser tudo aquilo livre de risco; protegido, acautelado, garantido, a própria palavra seguro já traz a idéia de proteção, prevenção, previdência. Outro conceito importante para entendimento é o fato gerador de todas as iniciativas previdenciárias existentes, que foi o risco aleatório, o risco, portanto, é à base do seguro. Existem riscos impossíveis de ser objeto de seguro, são chamados de riscos especulativos [6].

Apesar de não haver um consenso entre pesquisadores e escritores sobre o início do seguro comercial, pode se afirmar que o ponto chave foi a partir dos seguros marítimos, provavelmente entre florentinos e outros mercadores italianos no século XIII, nesta época o seguro marítimo alcançou um desenvolvimento grande, beneficiando cidades portuárias com coberturas securitárias [6].

O trabalho em questão terá o objetivo de desenvolver um *Data Warehouse*, baseado nos dados da Corretora de Seguros Karla, essa fundada em 1986 na cidade de Palmeira das Missões - RS, sendo que em 1995 mudou para cidade de Ijuí-RS, onde está localizada até a presente data, na Av 21 de Abril, 1062. A Corretora de Seguros Karla esta consolidada no mercado de corretagem de seguros, trabalhando, nos mais diversos ramos, com destaque para os ramos automóveis, residenciais, empresariais e seguros de vidas.





Modalidade do trabalho: Relato de experiência
Evento: 2011 SIT - I Seminário de Inovação e Tecnologia

Sabendo do crescimento tecnológico e suas vantagens, a empresa faz uso de um sistema de informações para o auxílio e controle da produtividade e clientes, e ainda de planilhas para outros controles, por esta razão, gera uma grande quantidade de dados armazenados, estes não são utilizados para auxílio a tomada de decisões, com essas informações, juntamente com a disponibilidade de acesso aos dados, nasceu a idéia de criar um DW para dar suporte à gestão de negócios. Nos próximos itens, descreveremos o desenvolvimento do DW aplicado aos dados da corretora de seguros.

Seleção dos Dados

A primeira impressão que temos ao realizar a movimentação dos dados para um de DW, é que será uma simples extração de dados entre um local e outro, porém, na prática não é bem assim que acontece, pois, é necessária a conversão dos dados e a criação da nova base de dados e após a integração destes na nova base [5].

A Corretora de Seguros Karla, utiliza para gerenciamento de seus negócios o sistema denominado Sistema de Gerenciamento para Corretora de Seguros (SGCS), este disponibilizado pela seguradora Bradesco para seus corretores. Tivemos contato também com outras fontes de dados oriundas de sistemas de cotação de seguradoras, planilhas, dentre outras, porém, após uma análise mais criteriosa decidimos não integrarmos estas no DW, levando em consideração que a conversão e utilização destas informações seriam demasiadamente difíceis, utilizando como parâmetro a relação custo x benefício.

Os dados selecionados para o desenvolvimento do *data warehouse* são dos últimos quatro anos, contidos no sistema SGCS, este período foi escolhido devido há algumas mudanças e migração de sistemas anteriores a essa data, esta decisão foi consensual com o administrador, sendo que, antes desse período não havia um controle muito rígido dos cadastros realizados e portanto esses dados poderiam não estar de acordo com a realidade e influenciar o resultado do trabalho de forma errônea. Na próxima sessão apresentaremos os métodos utilizados no processo de ETC.

Processo de ETC

Aqui, após a seleção dos dados que foram incluídos no DW, nos deteremos em apresentar os métodos que foram necessários para o processo de ETC, desde a conversão, critério de transformação e limpeza abordado, e ainda quais ferramentas foram utilizadas nesse processo.

Conversão e limpeza dos dados

Recebemos a base de dados no formato da extensão “*mdb*”, esse formato é originário do programa Microsoft *Access Office*, por isso, com o auxílio da ferramenta IBM DB2 foi possível a conversão para o formato com extensão “*csv*”, em seguida foi criada uma nova base relacional e suas respectivas tabelas, por último realizamos a carga destes dados nas novas tabelas, resultando em uma base relacional no DB2.

Após a criação da nova base dos dados, realizamos a limpeza dos dados, sendo que nesse ponto o maior ruído encontrado foi a falta de campos preenchidos, nos deparamos com vários campos sem nenhuma linha cadastrada, para solucionar este problema foi adotada a técnica de exclusão dos campos em branco, esta decisão foi tomada com o auxílio do



Modalidade do trabalho: Relato de experiência
Evento: 2011 SIT - I Seminário de Inovação e Tecnologia

especialista da área na análise das colunas retiradas, lembrando sempre que se for necessário a exclusão de alguma informação, é imprescindível ter a certeza que estes dados não irão comprometer o resultado.

Modelagem Multidimensional

A modelagem dimensional foi realizada neste momento, a tabela fato foi denominada Fato Venda de Seguros, e as dimensões: Tempo, Profissão, Cliente, Cidade, UF, Ramo do Seguro, Tipo do Seguro, Cia Seguradora, Apólice, Situação da Apólice, foram criadas com o auxílio do gestor da empresa, sendo essas as informações chave no ramo de negócios segurador desta corretora. A tabela fato centralizada, dimensões e métricas, essas definidas como custo de apólice, custo juros, custo IOF e premtotal.

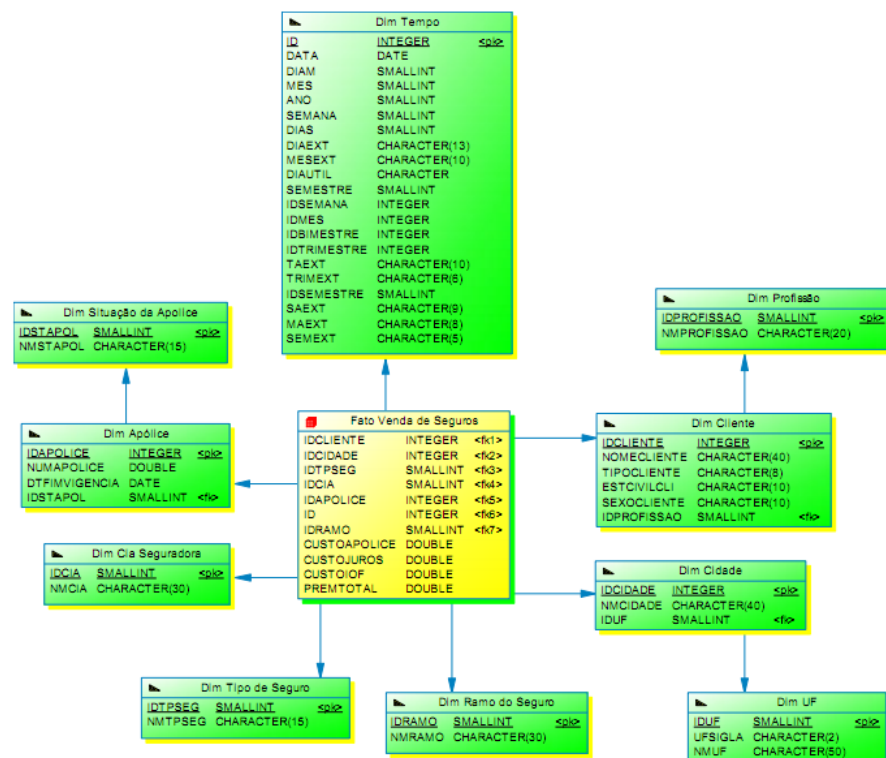


Figura 1. Modelagem Dimensional da Corretora de Seguros

Carga dos dados

A carga dos dados foi realizada após definida a modelagem, ou seja, a estrutura do *data warehouse*, neste ponto realizamos a criação das dimensões e da tabela fato, e em seguida realizamos o *load* dos dados através de *script* em SQL no IBM DB2 nas tabelas.

Visualização do DW

Por fim, após todo o processo de desenvolvimento utilizamos a ferramenta *Analysis Server* do *Microsoft Visual Studio 2005* para finalizar o DW, nela realizamos as ligações necessárias entre a tabela fato e suas dimensões para o funcionamento correto. Com isso, foi



Modalidade do trabalho: Relato de experiência
Evento: 2011 SIT - I Seminário de Inovação e Tecnologia

possível a visualização e funcionamento do DW, sendo que a grande facilidade de relacionar as informações entre as mais distintas dimensões foi o fator que mais impressionou o gerente. O nível de granularidade foi determinado pela dimensão tempo, que foi estruturada para atender os períodos, de ano, meses, semanas, e dias.

Conclusões

O *Data Warehouse* é um processo complexo, que demanda compreensão da área onde o projeto será realizado, por isso, a colaboração da empresa ou organização envolvida é de suma importância para que o resultado obtido seja satisfatório.

No desenvolvimento do DW observamos que todas as fases de desenvolvimento devem ser seguidas, porém, em determinados momentos pode ser necessário retornar a um ponto anterior e realizar ajustes necessários para o bom andamento do projeto.

Para finalizar, conseguimos entender a capacidade de suporte que o DW pode dar no auxílio a tomada de decisões, através de diversas fontes integradas, e a partir delas, dar a possibilidade de extrair informações, que em um ambiente normal de banco de dados não é possível.

Referências

- BALLARD C., Herreman D., Schau D., Bell R., Kim E., Valencia, Data Modeling Techniques for Data Warehousing, IBMRedBooks, IBM Corporation 1998.
- INMON, Willian H. Como construir o Data Warehouse; tradução de Ana Maria Netto Guz. Rio de Janeiro: Campus, 1997.
- Juswiak, André. Influenciadores na Decisão de Compra de Seguros – O Caso da Corretora de Seguros Karla Ltda. Trabalho de Conclusão de Curso, Unijuí. Ijuí, 2008.
- KIMBALL, Ralph. Data Warehouse Toolkit; tradução de Monica Rosemberg. São Paulo: Makron Books, 1998.
- NAISBITT, J., ABURDENE, P. (1990). Megatrends 2000. New York, USA. July, 1990.
- WARMELING, Kelvin Jacob. Protótipo de Data Warehouse aplicado a Cia de Seguros de Automoveis. Trabalho de Conclusão de Curso. Universidade Regional de Blumenau(URB). Blumenau/SC, 1999.