

Modelagem de uma BI (Business Intelligence) usando Software Livre: Caso de Uso Pentaho Community Edition

Alberto Teixeira e Silva
Tecnólogo em Sistema para Internet – Faculdade CET – albertothe@gmail.com.
Francisco Luciani de Miranda Vieira
Mestre em Informática – Faculdade CET – lucianivieira@gmail.com.

Resumo: Este Artigo científico apresentará o estudo e modelagem de uma *Business Intelligence* para auxiliar nas tomadas de decisões no ambiente gerencial de pequenas, médias e grandes empresas. O artigo mostrará a modelagem de um Data Warehouse, seu processo de ETL (*Extract, Transform and Load*), modelagem e criação de cubo OLAP (*Online Analytic Processing*), *Dashboard* e relatórios *Ad Hoc*.

Palavras chave: Pentaho, *Data WareHouse*, Cubos, *DashBoard*, Relatório.

1. Introdução

Com alta competitividade do mercado e com o grande volume de dados que a cada segundo cresce nas empresas, fica cada vez mais difícil sobreviver no mercado sem um setor exclusivo para inteligência de negócio, para que as tomadas de decisões estratégicas, além de precisas, sejam com a maior rapidez possível.

O Conceito de *Business Intelligence* (BI) resume-se em extrair dados muitas vezes de várias bases distintas, transformar e consolidar estas informações em conhecimento, para que sejam usadas de forma inteligente pelos gestores das empresas (Inomn, 2010).

Uma Business Intelligence quase sempre existe uma grande necessidade de se possuir um depósito de dados (Data Warehouse) para que os dados consolidados sejam armazenados e consultados posteriormente sem comprometer o desempenho do banco operacional. Um Data Warehouse é uma fonte de dados para consulta da organização (Kimball, 1998).

2. Procedimentos Metodológicos

Este Artigo mostrará a modelagem de uma BI com ferramentas da suíte *Pentaho Community Edition*, versão *Open Source* da empresa Pentaho, Abaixo segue uma breve apresentação dos componentes da mesma.

Pentaho User Console (PUC): é a plataforma integradora de todos os componentes da suíte.

Pentaho Data Integration (PDI): é um dos componentes mais poderosos da suíte, responsável pelo processo de ETL, mas que também tem uma infinidade de utilizações.

Pentaho Schema Workbanch (PSW): interface gráfica para criação de cubos OLAP.

Pentaho Report Designer (PRD): ferramenta da suíte para criação de relatórios com grande facilidade de manipulação e com muito profissionalismo no visual.

Community Dashboard Editor (CDE): excelente ferramenta para criação de Painéis de Indicadores (Dashboard), muito intuitivo e com uma grande variedade de templates.

O Artigo resume-se em controlar as informações de forma organizada, para que o resultado seja otimizado. Desta forma faremos análises criteriosas e um bom levantamento de requisitos dos dados da empresa, para que tenhamos um processo de ETL confiável.

Na construção do *Data Warehouse* (DW) utilizaremos a ferramenta PDI para o processo de ETL e a construção do armazém de dados utilizando um SGBD também *Open Source*, o PostgreSQL. Com o PDI faremos as transformações de cada dimensão do DW, onde todas as transformações serão executadas de forma automatizada por um único processo.

Com o DW construído, demonstraremos a criação do cubo OLAP, responsável pela visão multidimensional dos dados pelos gestores, utilizando a ferramenta PSW. O cubo OLAP também será usado para a criação das visões analíticas, utilizando a ferramenta SAIKU.

Os relatórios serão criados com a ferramenta PRD, que tem como principal característica a facilidade de uso, com alguns *templates* e deixando o desenvolvedor livre para criar os seus próprios relatórios. Para finalizar a BI utilizaremos o CDE para a criação dos Dashboard's, que são os painéis de indicadores.

3. Resultados e discussões

Com a modelagem completa os gestores da empresa em que a BI foi implantada terão um ganho significativo na tomada de decisão e aumentarão sua capacidade de competitividade no mercado. Teremos ganho de desempenho no SGBD, pois nosso *Data Warehouse* foi criado de forma desnormalizada e as consultas mais complexa são realizadas diretamente no mesmo.

Uma desvantagem notada é com relação ao período dos dados, já que os dados operacionais estão em outros armazém de dados, que ficam atualizando em tempo real, diferente do Data Warehouse que é atualizado de acordo com suas diretrizes, geralmente usado D-1 (um dia antes do real).

4. Considerações finais

Em resumo, implantação de uma *Business Intelligence* é de grande utilidade não só em grandes corporações, mas também em ambientes de pequenas e médias empresa, já que vemos que a competitividade aumenta a cada dia. Notamos que o resultado final (interação BI - Gestores) da modelagem representa muito pouco do trabalho geral da implantação, já que a parte de levantamento de requisitos e processos de ETL leva boa parte do tempo.

5. Referências

- CASTERS, M. **Pentaho.3.2.Data.Integration**.Beginners Guide, 2010.
- DATE, C. J. **Introdução a Sistemas de Banco de Dados**, Volume 1, Tradução da 8ª edição americana, Editora Campus, 2004.
- INMON, W.; TERDERMAN, R.; IMHOFF, C. - **Data Warehousing: como transformar informações em oportunidades de negócios**. Editora Berkely, 2010.
- KIMBALL, R. **Data Warehouse tool kit: técnicas para construção de data warehouses dimensionais**. Makron Books, 1998.
- PULVIRENTI, A. S.; ROLDÁN, M. C. **Pentaho Data Integration 4 Cookbook**, 2011.