



INSTITUTO FEDERAL DE
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA
BAHIA
Campus Eunápolis



pindorama
Revista Eletrônica Científica do IFBA

Revista Eletrônica Multidisciplinar Pindorama do Instituto Federal de
Educação, Ciência e Tecnologia da Bahia – IFBA Nº 02 – Ano 3 – junho/2012 –
www.revistapindorama.ifba.edu.br

UM ESTUDO SOBRE A VIABILIDADE DE IMPLANTAÇÃO DE UM *DATA WAREHOUSE* EM INSTITUTOS FEDERAIS DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA

LAÉRCIO IVES SANTOS

ESPECIALISTA EM BANCO DE DADOS PROFESSOR DO ENSINO BÁSICO
TÉCNICO E TECNOLÓGICO INSTITUTO FEDERAL DE EDUCAÇÃO,
CIÊNCIA E TECNOLOGIA DO NORTE DE MINAS GERAIS, ARAÇUAÍ-MG
LATTES: <http://lattes.cnpq.br/5234224884543563> E-MAIL:
laercio.santos@ifnmg.edu.br

RESUMO

O trabalho propõem uma análise da viabilidade de se implantar um *Data Warehouse*(DW) em Institutos Federais de Educação, Ciência e Tecnologia (IFs) tomando como estudo de caso o Instituto Federal de Educação, Ciência e Tecnologia do Norte de Minas Gerais - IFNMG. Assim, a tecnologia de DW é apresentada, abordando características, arquiteturas, e formas de implementação.

Palavras chaves: *Data Warehouse*, Tomada de decisão, Institutos Federais de Educação, ciência e tecnologia.

Introdução

A ideia de que a informação é o maior bem de uma organização, consegue mais seguidores com o passar do tempo. É notável ainda que, uma

grande parte dessas informações está disponível nas próprias organizações. Os dados para a obtenção de informações estão armazenados nos bancos de dados operacionais, sistemas legados, planilhas, arquivos pessoais e documentos. O grande desafio é como torná-los úteis, transformando dados em informações e, estas, em ações que melhorem a gestão dos negócios. Como acessar essa imensa base de dados dispersa e concretizá-la em uma fonte de informação viável? O objetivo é extremamente simples: transformar e organizar os dados para geração de informações estratégicas úteis.

Uma importante questão estratégica para o sucesso de qualquer organização nos dias atuais é a sua capacidade de analisar, planejar e reagir, rápida e imediatamente às mudanças nas condições de seus negócios. Para que isso aconteça, é necessário que a organização disponha de mais e melhores informações, que constituem, reconhecidamente, a base destes processos (CAMPOS; ROCHA FILHO, 2000 apud CLEMES, 2001, p.15).

Com base nisso, o trabalho em questão tem como objetivo principal elaborar um estudo de viabilidade de implantação de uma arquitetura de *Data Warehouse*(DW) no Instituto Federal de Ciência, Educação Tecnologia do Norte de Minas Gerais. Esse trabalho foi baseado em (CLEMES, 2001), em que o autor propõe uma arquitetura de DW em instituições de ensino superior, afim de propiciar aos administradores acesso a informações institucionais em tempo real em um único ambiente e a partir de qualquer computador conectado na internet.

O Instituto Federal do Norte de Minas Gerais foi criado a partir da parceria entre o antigo CEFET Januária e a Escola Agrotécnica Federal de Salinas que se transformaram nos campi Januária e Salinas. Entretanto, foram criados cinco novos campi, alguns ainda em fase de implantação, eles serão instalados nas cidades de: Montes Claros, Pirapora, Arinos, Araçuaí e Almenara. Assim, não existe um mecanismo automatizado que auxilie no processo de tomada de decisão por parte dos diretores dos campi bem como da reitoria.

O trabalho foi desenvolvido obedecendo às seguintes etapas;

- ✓ Levantamento bibliográfico das tecnologias existentes no mercado e na literatura.
- ✓ Levantamento de sistemas legados bem como de fontes de informações contidas no formato não digitalizado nos campi Januária-MG e Salinas-MG.
- ✓ Desenvolvimento do estudo proposto.

1. Data Wharehouse

A tecnologia *Data Warehouse (DW)* ou Armazém de Dados surgiu com a necessidade de um ambiente integrado, onde fossem armazenados e gerenciados eficientemente os dados produzidos pelas aplicações operacionais, aplicações que não são capazes de gerar conhecimento para uma organização.

Podemos encontrar algumas definições para o termo, assim, em Almeida (2004) um *DW* é definido como um banco de dados que contem o histórico de uma organização, onde as diversas áreas podem planejar suas estratégias através dos dados contidos nesses bancos de dados.

Ainda segundo Almeida, em geral um *DW* é utilizado pelos setores administrativos e pela alta gerência de uma organização, no desenvolvimento de análises e relatórios. Os dados inseridos nesse *DW* foram previamente consolidados, sumarizados e depurados para serem utilizados num ambiente orientado a consultas com uma modelagem específica para esta finalidade.

Todos os dias, as organizações geram e captam dados sobre todos os aspectos dos seus negócios. Na grande maioria das vezes, estes dados estão sendo armazenados e manipulados para suportar as atividades críticas (faturamento, almoxarifado, contabilidade, entre outros) através dos sistemas de processamento de transações *on-line* (OLTP), sistemas esses que têm como característica principal o baixo tempo de resposta para processar as suas requisições.

INMON(1997) destaca ainda a descrença dos dados, dizendo que, a distância entre as informações provenientes de dois departamentos que usam

sistemas não integrados, pode levar a uma tomada de decisão baseada em interesses políticos ou opiniões pessoais.

Pode-se notar que decisões baseadas em ferramentas *OLTP*, além de não ser confiáveis, são demoradas. E ainda não atendem de maneira correta aos gestores de uma organização muito dinâmica.

1.1. *Data warehouse* como ferramenta de apoio a tomada de decisão em instituições de ensino

Em CLEMES (2001), pode-se encontrar um trabalho voltado para a arquitetura e implantação de um *DW* em instituições do ensino superior. Nesse trabalho o autor enfatiza o uso pioneiro desse tipo de tecnologia em instituições de ensino. Um *DW* geralmente é utilizado por organizações com fins lucrativos e que funcionam baseadas na competitividade.

O trabalho propõe uma arquitetura de *DW* para Instituições de ensino superior usando como estudo de caso uma implementação na Universidade Federal de Santa Catarina - UFSC. Assim, essa arquitetura é implantada e testada.

O objetivo de se implantar um *DW* em um IES é propiciar aos gestores da instituição uma economia de tempo e esforço no processo de tomada de decisão bem como a facilidade de responder auditorias e comprovar indicadores da instituição(CLEMES,2001).

2. Institutos federais de educação, ciência e tecnologia - IF

Os Institutos Federais são instituições de educação superior, básica e profissional, pluricurriculares e multicampi, especializados na oferta de educação profissional e tecnológica nas diferentes modalidades de ensino, com base na conjugação de conhecimentos técnicos e tecnológicos com as suas práticas pedagógicas(BRASIL, 2008).

Para efeito da incidência das disposições que regem a regulação, avaliação e supervisão das instituições e dos cursos de educação superior, os Institutos Federais são equiparados às universidades federais.

A Lei 11892 de 22 de Dezembro de 2008 criou 38 Institutos Federais a partir da junção das Escolas Agrotécnicas Federais e dos Centros Federais de Educação Tecnológica - CEFETs. Assim, o Instituto Federal de Educação, Ciência e Tecnologia do Norte de Minas Gerais, foi criado fundindo a Escola agrotécnica Federal de Salinas e o Centro Federal de Educação Tecnológica de Januária-MG.

Os Institutos Federais tem como finalidades e características [BRASIL, 2008]:

- ✓ Ofertar educação profissional e tecnológica, em todos os seus níveis e modalidades, formando e qualificando cidadãos com vistas na atuação profissional nos diversos setores da economia, com ênfase no desenvolvimento socioeconômico local, regional e nacional;
- ✓ Desenvolver a educação profissional e tecnológica como processo educativo e investigativo de geração e adaptação de soluções técnicas e tecnológicas às demandas sociais e peculiaridades regionais;
- ✓ Promover a integração e a verticalização da educação básica à educação profissional e educação superior, otimizando a infraestrutura física, os quadros de pessoal e os recursos de gestão;
- ✓ Orientar sua oferta formativa em benefício da consolidação e fortalecimento dos arranjos produtivos, sociais e culturais locais, identificados com base no mapeamento das potencialidades de desenvolvimento socioeconômico e cultural no âmbito de atuação do Instituto Federal;
- ✓ Constituir-se em centro de excelência na oferta do ensino de ciências, em geral, e de ciências aplicadas, em particular, estimulando o desenvolvimento de espírito crítico, voltado à investigação empírica;
- ✓ Qualificar-se como centro de referência no apoio à oferta do ensino de ciências nas instituições públicas de ensino, oferecendo capacitação técnica e atualização pedagógica aos docentes das redes públicas de ensino;
- ✓ Desenvolver programas de extensão e de divulgação científica e tecnológica;

- ✓ Realizar e estimular a pesquisa aplicada, a produção cultural, o empreendedorismo, o cooperativismo e o desenvolvimento científico e tecnológico;
- ✓ Promover a produção, o desenvolvimento e a transferência de tecnologias sociais, notadamente as voltadas à preservação do meio ambiente.

2.1. O processo de tomada de decisão dentro de um IF

O processo de tomada de decisão em um Instituto Federal (IF) envolve muitos aspectos. Um desses aspectos está relacionado à democracia, tendo como princípio básico a participação ativa nas decisões de maior relevância de todos os membros e participantes do ambiente escolar. Não só as autoridades de um Instituto federal como Reitor, Pro – Reitores, Diretores, mas também professores, alunos, pais e a sociedade onde aquela Instituição está inserida.

O órgão principal de um IF é o conselho superior, este é formado pelo Reitor, professores, técnicos administrativos, alunos e ex-alunos. Assim, o conselho superior tem como principal função tomar as principais decisões no âmbito de um IF.

No estatuto do IFMG são definidas as principais funções do conselho superior, que são:

- ✓ Aprovar as diretrizes para atuação do IFMG e zela pela execução de sua política educacional.
- ✓ Aprovar as normas e coordena o processo de consulta à comunidade escolar para escolha do Reitor do IFMG e dos Diretores Gerais dos Campi.
- ✓ Aprovar os planos de desenvolvimento institucional (PDI) e de ação e aprecia a proposta.
- ✓ Aprovar o projeto político-pedagógico, a organização didática, regulamentos internos e normas disciplinares.
- ✓ Aprovar normas relativas à acreditação e à certificação de competências profissionais, nos termos da legislação vigente.

- ✓ Autorizar o Reitor a conferir títulos de mérito acadêmico e outras honrarias.
- ✓ Apreciar as contas do exercício financeiro e o relatório de gestão anual, emitindo parecer conclusivo sobre a propriedade e regularidade dos registros.
- ✓ Deliberar sobre taxas, emolumentos e contribuições por prestação de serviços em geral a serem cobrados pelo IFNMG.
- ✓ Autorizar a criação, alteração curricular e extinção de cursos no âmbito do IFNMG, bem como o registro de diplomas.
- ✓ Aprovar a estrutura administrativa e o regimento geral do IFNMG.
- ✓ Deliberar sobre questões submetidas a sua apreciação.

2.2. Sistemas de informações em instituto federal

No IFNMG existem três tipos de sistemas de informação, o sistema de gestão administrativa, o sistema de gestão acadêmica e o sistema de gestão de bibliotecas.

O sistema de gestão de Bibliotecas está em fase de aquisição pelo processo de licitação. O edital do processo de licitação está sendo construído pelos bibliotecários do IFNMG para apreciação posterior dos gestores da instituição.

O sistema Siga-ADM está sendo elaborado pela RENAPI (Rede de Pesquisa e Inovação em Tecnologias Digitais) sendo um subsistema do projeto Siga-EPT (Sistema Integrado Gestão Acadêmica). O projeto visa à elaboração de um sistema integrado de gestão acadêmica desenvolvido com tecnologias livres e de forma colaborativa por várias instituições federais do Brasil. O projeto tem o apoio do Ministério da Educação do Brasil - MEC, através da Secretaria da Educação Profissional e Tecnológica – SETEC (RENAPI, 2010).

Os módulos administrativos contemplados pelo SIGA-ADM oferecem as seguintes funcionalidades: Requisições online de Almoxarifado, Compras, Diárias, Hotel, Passagens; Restaurante, Serviços externos e Veículos; Controle de pagamento de bolsas oferecidas pela instituição; Licitação de materiais e

serviços; Controle orçamentário; Controle contábil; Controle de empenhos; Controle de pagamentos; Patrimônio; Almoxarifado; Protocolo e Recursos Humanos.

Atualmente existem dois sistemas de informação para auxílio nos trabalhos acadêmicos no IFNMG. Um sistema está implantado no campus Januária-MG e outro está no campus Salinas-MG. São eles:

- ✓ O AIX GIZ é um sistema de controle acadêmico desenvolvido na linguagem de programação DELPHI e utiliza o Sistema Gerenciador de Banco de Dados MS *SQL Server*. Esse sistema está em uso no campus Januária-MG;
- ✓ O Sala é um sistema de controle acadêmico desenvolvido na linguagem de programação *Clipper*, utiliza o Sistema Gerenciador de Banco de Dados *DBase*. Esse sistema está em uso no campus Salinas-MG;

Entretanto, o IFNMG já está em processo de aquisição de um sistema integrado de gestão acadêmica. A comissão de Tecnologia da Informação do IFNMG está trabalhando para montar um processo de licitação para aquisição desse sistema. Assim, os dados contendo nesses dois sistemas seriam migrados para o sistema integrado adquirido através do processo de licitação.

3. Principais custos na implantação de um *data warehouse*

A implantação de um *DW* exige um investimento. Esse investimento pode ser dividido em duas etapas: custos iniciais e custos periódicos [(NMON, 1998). Ainda existem os custos do ponto de vista do tipo de investimento, que pode ser dividido em custos capitais e custos operacionais. Os custos capitais são decorrentes do gasto de dinheiro referente a compra de equipamentos, serviços e estrutura. Os custos operacionais é o gasto com pessoal, não necessariamente gastos financeiros, mas sim gastos de tempo para o desenvolvimento de soluções. Este tempo é precioso e não é somente gastando-se mais dinheiro com pessoal que o tempo de desenvolvimento do *data warehouse* irá diminuir.

Os custos iniciais dizem respeito aos custos no ponto de partida do *DW*. Já os custos periódicos estão relacionados aos custos para manutenção da plataforma do *DW*. Assim, os custos iniciais podem ser classificados como:

Hardware

- ✓ Processamento
- ✓ Armazenamento
- ✓ Estrutura de rede

Software

- ✓ Sistema de Gerenciamento de Banco de Dados(SGBD)
- ✓ Ferramentas de gerenciamento e monitoramento do sistema
- ✓ Criação e povoamento dos metadados
- ✓ Criação ou aquisição dos programas de ETL (*Extraction, Transform and Load*)

Pessoal

- ✓ Consultoria em *DW*
- ✓ Administrador de *data warehouse* (banco de dados)
- ✓ Administrador de rede
- ✓ Analistas de sistema
- ✓ Programadores
- ✓ Treinamento dos usuários

O custo com *hardware* diz respeito ao investimento na estrutura física de suporte ao *DW*, como computadores, discos rígidos, dispositivos de fitas de dados, infraestrutura de rede. Assim, o maior gasto nesse quesito é a parte de armazenamento, pois, um grande *DW* pode chegar a ocupar vários *Tera Bytes* de dados. Então, para diminuir esse gasto, são instalados sistemas de armazenamento em massa, como fitas magnéticas e armazenamento ótico, diminuindo o custo muitas vezes. Deve-se realizar uma previsão do tamanho inicial do *DW* e uma estimativa de crescimento do *data warehouse* para que seja adquirido o *hardware* necessário para conter o crescimento.

Os gastos com pessoal são de extrema importância para a implantação de um *DW*. Inicialmente, deve-se contratar uma consultoria para uma avaliação

das necessidades e proposição das melhores alternativas para a criação do *DW*. A partir daí, contratam-se ou se realocam profissionais de administração do *DW*, redes, analistas e programadores.

Nesse custo com pessoal vale destacar o treinamento dos usuários finais. Em um sistema de informação, formado por software, hardware e pessoas, um grande problema está na capacitação do usuário. Muitas das vezes o usuário não tem conhecimento de como usar corretamente o sistema tornando-o inviável para a organização.

Os custos periódicos se dividem em:

- ✓ Manutenção de *hardware*;
- ✓ Manutenção do SGBD
- ✓ Manutenção das ferramentas de ETL
- ✓ Aquisição de novas ferramentas
- ✓ Manutenção de pessoal: analistas, programadores, Administradores de base de dados.

Os gastos periódicos com *hardware* estão relacionados com manutenção do equipamento existente, devido ao desgaste natural das peças, e com acréscimo de dispositivos de processamento/armazenamento de acordo com o crescimento do *DW*

Os gastos com software estão em sua maioria relacionados com o gasto de pessoal, como a manutenção do software e da estrutura de metadados. Eventualmente, pode haver aquisição de novas ferramentas para a manutenção e uso do *DW*.

4. Principais benefícios com a implantação de um *data warehouse*

Um dos principais benefícios que um *DW* pode trazer para uma organização o acesso às informações de forma rápida e eficiente. É visível a facilidade de extração de relatórios.

Ao extrair relatórios sem um *DW*, uma organização deveria seguir algumas etapas como: verificar em quais sistemas e base de dados estão os

dados, padronizar os dados, conciliar os dados, montar a estrutura para os dados, reunir os dados e por fim construir o relatório.

A força de desenvolvimento envolvida em produzir um relatório é muito custosa. E não podemos reaproveitar relatórios, pois as necessidades de dados mudam para cada relatório. Assim, utilizando um *DW* para se construir um relatório, o processo se simplifica, e algumas etapas podem ser descartadas indo direto ao passo de construir o relatório. Desse modo a construção do relatório terá um prazo mais curto.

INMON(1997) divide os usuários de um *DW* entre “Fazendeiro e Exploradores” em que os fazendeiros são os usuários que utilizam o *DW* no dia a dia enquanto os exploradores se classificam como usuários que não usam o *DW* no cotidiano. Porém, em uma dessas “explorações” podem encontrar uma pedra preciosa em forma de informação. Assim, os custos de implantação de um *DW* se justificam na ótica de um fazendeiro, um usuário que utiliza as ferramentas do *DW* no cotidiano.

Um benefício na implantação de um *DW* em um Instituto Federal é propiciar

aos administradores da instituição uma economia de tempo e esforço no processo

de tomada de decisão. Outra vantagem que se deve evidenciar é a facilidade em responder auditorias e comprovar os indicadores da instituição.

Vale destacar algumas questões que podem ser respondidas com o auxílio de uma estrutura de *DW* em um Instituto federal:

- ✓ Perfil dos alunos ingressos com base nos dados das provas de seleção e das matrículas;
- ✓ Perfil dos alunos em relação aos empréstimos de livros com base no histórico de empréstimos;
- ✓ Perfil das compras com base no histórico e nos dados dessas compras;
- ✓ Perfil de cada disciplina, cursos e professores;
- ✓ Perfil de contas a pagar com base no histórico dessas;

Dessa forma um *DW* em um Instituto Federal pode auxiliar na tomada de decisões como, abertura de novos cursos, aquisição de livros, abertura de turmas, aquisição de materiais de uma maneira geral, formas para reduzir gastos, metodologias para melhoria do ensino, metodologias para melhoria de cursos isolados ou disciplinas isoladas, dentre outras. E auxiliar em possíveis processos de auditorias promovendo o acesso a informações de uma forma menos custosa.

Conclusão

É fato que uma tecnologia de *Data Warehouse* pode trazer muitos benefícios para uma organização. Esse tipo de tecnologia auxilia no processo de tomada de decisões por parte dos gestores dessa organização. Também é fato que as aplicações de *Data Warehouse* são usadas de forma mais ampla em organizações que tem fins lucrativos, organizações que tem como maior objetivo atingir o maior número de clientes possível.

A tecnologia de *Data Warehouse* em organizações sem fins lucrativos, como organizações públicas de ensino, tem como principal objetivo reduzir o custo de obtenção de informação. Assim, acelerando o processo de tomada de decisão e auxiliando no processo de auditoria quando for necessário. A tecnologia por si só não vai melhorar os processos de uma organização, porem, um *Data Warehouse* tem como principal objetivo servir de suporte para os gestores da organização.

No estudo feito, são mostrados os principais custos em se implantar um *Data Warehouse*, ainda que não exista na literatura uma forma de mensurar esses custos. Assim, é notado que deveria ser feito um estudo mais detalhado abordando as dimensões de todos esses custos.

Um *Data Warehouse* é mais indicado para organizações que tenham base de dados utilizadas ao longo de muitos anos e que chegam a aproximadamente 100 Giga Bytes. Assim, em uma organização que ainda está

começando a implantar os sistemas de informação talvez não seja viável a implantação desse tipo de tecnologia.

Essa implantação traria benefícios para a tomada de decisão da organização, porém, o custo inicial é alto e a percepção dos resultados poderia ser demorada. Dessa forma, esse tipo de tecnologia teria maiores resultados se implantada após algum tempo de implantação dos sistemas de informação. Visto que a estrutura interna da organização ainda não está totalmente montada.

Como trabalhos futuros podemos citar o projeto e implementação do *Data Warehouse*, assim a partir desse estudo de viabilidade, se começaria um projeto levantando os fatos, as dimensões e arquitetura como um todo para se implantar essa tecnologia em um Instituto Federal.

Podemos citar também um mecanismo de Mineração de Dados para descoberta de tendências dentro dessa base de dados do Instituto Federal.

Referências bibliográficas

ALMEIDA, Eduardo Cunha. **Estudo de viabilidade de uma plataforma de baixo custo para *data warehouse***. Curitiba, 2004. Disponível EM http://dspace.c3sl.ufpr.br/dspace/bitstream/1884/662/1/Eduardo_Cunha_de_Almeida.pdf>. Acesso em: 20 de Maio de 2010.

BRASIL, 11892 de 22 de Dezembro de 2008.

CAMPOS, M. L.; ROCHA FILHO, A. V. ***Data warehouse***. Disponível em:

<<http://genesis.nce.ufrj.br/dataware/tutoria/home.html>>. Acesso em: 08 jun. 2010.

CLEMES, M. ***Data Warehouse como suporte ao sistema de informações gerenciais em uma instituição de ensino superior: estudo de caso na UFSC***. Florianópolis, 2001.

INMON, W. H. ***Como construir o Data Warehouse***. 2ª ed.- Rio de Janeiro:

Campus, 1997.

RENAPI. 2010. **Rede de Pesquisa e Inovação em Tecnologias Digitais**. Disponível em: <http://www.renapi.org/sigaept/o-projeto/sistemas>. Acesso: 19/06/2010.