



**UNIVERSIDADE ESTADUAL DO MARANHÃO
PRO-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
CENTRO DE CIÊNCIAS TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA
DE COMPUTAÇÃO E SISTEMAS**

ANTONIO VALBER LIMA SANTANA

**UMA ABORDAGEM BASEADA NO PROCESSO DE ETL INTEGRADO AOS
INDICADORES DE NEGÓCIOS ATRAVÉS DE VISÕES MATERIALIZADAS
EM DATA WAREHOUSE PARA AUXILIAR NA TOMADA DE DECISÃO**

DISSERTAÇÃO DE MESTRADO

SÃO LUIS

2016

ANTONIO VALBER LIMA SANTANA

**UMA ABORDAGEM BASEADA NO PROCESSO DE ETL INTEGRADO AOS
INDICADORES DE NEGÓCIOS ATRAVÉS DE VISÕES MATERIALIZADAS
EM DATA WAREHOUSE PARA AUXILIAR NA TOMADA DE DECISÃO**

Dissertação apresentada ao Programa de Mestrado Profissional em Engenharia de Computação e Sistema, Departamento de Engenharia de Computação, Universidade Estadual do Maranhão, como requisito para a obtenção do título de Mestre em Engenharia de Computação e Sistemas. Área de Concentração: Tecnologia da Informação: Linha de Pesquisa Banco de Dados

Orientador: Prof. Dr. João Carlos Pinheiro
Coordenador: Prof. Ms. Henrique Mariano Costa
Amaral

**SÃO LUIS
2016**

**UNIVERSIDADE ESTADUAL DO MARANHÃO
PRO-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
CENTRO DE CIÊNCIAS TECNOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO E
SISTEMAS**

**UMA ABORDAGEM BASEADA NO PROCESSO DE ETL INTEGRADO AOS
INDICADORES DE NEGÓCIOS ATRAVÉS DE VISÕES MATERIALIZADAS
EM DATA WAREHOUSE PARA AUXILIAR NA TOMADA DE DECISÃO**

ANTONIO VALBER LIMA SANTANA

Aprovado em: 31/03/2016

BANCA EXAMINADORA:

**Prof. Dr João Carlos Pinheiro (IFMA)
(Presidente/Orientador)**

**Profa. Dra. Jeane Silva Ferreira (IFMA)
Examinador Externo**

**Prof. Msc. Josenildo Costa da Silva (IFMA)
Examinador Interno**

SÃO LUÍS

2016

Santana, Antônio Valber Lima.

Uma abordagem baseada no processo de ETL integrado aos indicadores de negócios através de visões materializadas em data warehouse para auxiliar na tomada de decisão / Antônio Valber Lima Santana. – São Luís, 2016.

64 f

Dissertação (Mestrado) – Curso de Engenharia de Computação e Sistemas, Universidade Estadual do Maranhão, 2016.

Orientador: Prof. Dr. João Carlos Pinheiro.

1. Indicadores de negócios. 2. Processo de ETL. 3. Visões materializadas. 4. Data warehouse. I. Título

CDU: 004.62:005.53

Dedico esta pesquisa

A minha mãe,
Maria dos Anjos Lima de Almada.

Aos meus filhos
João Pedro, Isabela Beatriz,
Lara Luciana e Antonio Valber.

A minha esposa Ruth Máximo.

A meus irmãos, Francisco Walter (*in memoriam*) e
Vagner José (*in memoriam*), sobrinhos,
amigos, colegas de curso, professores e
funcionários da UEMA e IFMA.

AGRADECIMENTOS

Agradeço a Deus, pela minha vida, saúde, sabedoria e oportunidades;

Aos professores e colaboradores da UEMA e IFMA

Ao meu orientador, professor Dr. João Carlos Pinheiro, pelos conselhos, dedicação e ter proporcionados oportunidades para ampliação de meus conhecimentos;

Aos meus professores do mestrado Josenildo Costa da Silva, Luís Carlos Costa Fonseca, Eveline de Jesus Viana Sá, João Belo Salgado, Rogério Moreira Lima Silva, Ewaldo Eder Carvalho Santana, Fernando Demétrio e ao coordenador do mestrado, Henrique Mariano Costa do Amaral;

A professora Dra. Jeane Silva Ferreira (IFMA), pelos conselhos e por todas as contribuições durante o desenvolvimento desta pesquisa;

Aos colegas do mestrado

Pelos bons momentos compartilhados durante o mestrado: Lanyllo Araújo, Francisco da Conceição, Idovaldo Cunha, Marcos Serra, Marcos Santos, Marcos de Jesus, José Linhares, Elisangela Cerqueira, Leão Júnior, Nonilton Alves, Emmanuel Xavier, Carlindo Lisboa, Jardiel Almeida, Marcio Fialho, José Antonio Moura, Marlon Bruno e Elda Sena;

A minha família

A minha mãe Maria dos Anjos e minha tia Maria Benedita que compartilham comigo, o sonho de uma vida próspera;

Ao meu pai Milton José Santana (*in memoriam*), aos meus irmãos Francisco Valter (*in memoriam*), Vagner José (*in memoriam*), Tania Maria, Joana Darc e Jorge Voney, que representam a minha família e são pessoas importante nesta minha vida;

Aos meus filhos João Pedro, Isabela Beatriz, Lara Luciana e Antonio Valber, que são o meu norte nesta jornada da vida;

A minha esposa Ruth Máximo, por todo seu carinho, amor e paciência nesse período do mestrado;

Aos meus primos, sobrinhos, tios pelo companheirismo e força para a realização de mais essa etapa;

A Escola Crescimento

Na figura de sua diretora geral Sra. Ariadine Bacelar, pelo compartilhamento da base de dados corporativa da empresa e seus indicadores de negócios;

Aos colegas de trabalho André Camacho, Luzivan Moraes e Patrícia Fonseca que contribuem para meu crescimento profissional.

Aos amigos

Lanyldo Araújo, Francisco da Conceição e Idovaldo Cunha, pelas inúmeras reuniões durante o mestrado, pela contribuição no desenvolvimento de artigos científicos e desta dissertação;

As professoras Janice Almada, Luciana Batista, Mariangela Dourado, Triciane Rabelo e Jhonatan Almada pelas correções em artigos científicos e na dissertação de mestrado;

Aos amigos e professores Gesiel Rios, Gilliard Leda e Ricardo Queiroz pela confiança depositada e registrada nas cartas de recomendação entregues na fase de seleção do mestrado;

E a todos que, de alguma forma contribuíram para que eu pudesse concluir esse trabalho.

*Só um ser perfeito pode dar a um indivíduo
imperfeito a ideia dele mesmo.*

(René Descartes)

RESUMO

SANTANA, Antonio Valber Lima. UMA ABORDAGEM BASEADA NO PROCESSO DE ETL INTEGRADO AOS INDICADORES DE NEGÓCIOS ATRAVÉS DE VISÕES MATERIALIZADAS EM DATA WAREHOUSE PARA AUXILIAR NA TOMADA DE DECISÃO. 64 f. Dissertação – Programa de Pós-Graduação em Engenharia de Computação e Sistemas, Universidade Estadual do Maranhão, São Luís, 2016.

Apesar das tecnologias e inovações no mercado ainda é uma prática comum em pequenas e médias empresas, a obtenção manual dos indicadores de negócios. Neste trabalho propõe-se uma abordagem para que o processo de extração, transformação e carregamento de dados (*Extract Transform Load* - ETL) seja integrado aos indicadores de negócios (*Key Performance Indicator* - KPI), através de visões materializadas. Este processo de integração tem as seguintes vantagens: (i) Ao invés de replicar a base de dados operacional por completo, somente transportará os dados correspondentes aos indicadores de negócios. Dessa forma pode-se além do ganho de performance dos dados, reduzir o espaço alocado em disco. (ii) Com a mudança da base de dados, a atualização no *Data Warehouse* ocorre apenas se as tabelas forem pertinentes aos indicadores de negócio e, conseqüentemente, temos uma atualização mais rápida. (iii) Com o uso de visões materializadas para otimização das consultas, os dados ficam disponibilizados quase que em tempo real.

Palavras Chaves: Indicadores de negócios, Processo de extração, transformação e carregamento de dados, Visões Materializadas e Data Warehouse

ABSTRACT

SANTANA, Antonio Valber Lima. AN APPROACH BASED ON INTEGRATED ETL PROCESS TO BUSINESS INDICATORS THROUGH MATERIALIZED VIEWS IN DATA WAREHOUSE TO ASSIST IN DECISION-MAKING.62 f. Dissertação – Programa de Pós-Graduação em Engenharia de Computação e Sistemas, Universidade Estadual do Maranhão, São Luís, 2016.

Although the technologies and innovations in the market is still a common practice in small and medium enterprises, manual obtain business indicators. In this research we propose an approach to the process of extraction, transformation and loading of data (Extract Transform Load -ETL) is integrated with business indicators (Key Performance Indicator - KPI) through materialized views. This process of integration has the following advantages: (i) Instead of replicating the operational database completely, only carry the data for business indicators. Thus it can be in addition to the performance data of the gain, reduce the space allotted on disk. (li) With the change of the database, updating the data warehouse would occur only if the tables were relevant to the business indicators and, consequently, a faster refresh. (lii) With the use of materialized views for optimizing queries, the data is available in near real time.

Key Words: Key Performance Indicator, Extract Transform Load, View Materialized and Data Warehouse

SUMÁRIO

Capítulo 1 - INTRODUÇÃO	14
1.2 Objetivos	16
1.3 Formulação do Problema	16
1.4 Organização da Dissertação	18
Capítulo 2 – FUNDAMENTAÇÃO TEÓRICA.....	19
2.1 Indicadores Chaves de Desempenho.....	19
2.2 Processo ETL.....	21
2.3 Data Warehouse	25
2.4 Visão Materializada	30
2.5 Trabalhos Relacionados.....	32
Capítulo 3 – ABORDAGEM PROPOSTA	34
3.1 Etapas da Pesquisa	34
3.2 Estudo de Caso.....	35
3.3 Arquitetura do processo de ETL atual	36
3.4 A abordagem proposta.....	37
3.5 Processo de ETL integrado aos indicadores de negócios	38
3.6 Atualização das Tabelas no <i>Data Warehouse</i>	40
3.7 Visões materializadas modeladas em indicadores de negócios em aplicações de Data Warehouse.....	41
Capítulo 4 – ANÁLISE DOS RESULTADOS	43
4.1 Ambiente de execução	43
4.2 Análises dos resultados.....	43
4.3 Experimentos	44
4.3.1 <i>Experimento 1</i>	44
4.3.2 <i>Experimento 2</i>	45
4.3.3 <i>Experimento 3</i>	45
Capítulo 5 – CONCLUSÃO	47
REFERÊNCIAS	49

LISTA DE FIGURAS

Figura 1- Cenário carregamento de dados - Arquitetura Atual Erro! Indicador não definido.	
Figura 2- Modelo de dados - KPI Efetividade de Compras Erro! Indicador não definido.	
Figura 3- Modelo de Dados – Visão Materializada Ambiente de Data Warehouse Erro! Indicador não definido.	
Figura 4- Organização da Dissertação Erro! Indicador não definido.	
Figura 5- Processo de ETL, (Brave River Solutions)..... Erro! Indicador não definido.	
Figura 6- Fases de implantação do processo de ETL.....	23
Figura 7- Principais componentes, adaptado Turban et al, 2009.....	28
Figura 8- Arquitetura do Data Warehouse (adaptado: Primak, 2008)	29
Figura 9- Consulta VC e Visão VMC	31
Figura 10- Arquitetura atual do ETL e Aplicação Data Warehouse (Adaptado: Rob, 2011)	37
Figura 11- KPI integrado ao processo ETL	37
Figura 12- Processo de ETL modelados em indicadores de negócios utilizando visões	38
Figura 13- Visão V_KPI1	39
Figura 14- Processo de Carga Inicial.....	39
Figura 15- Modelo da Tabela DW_LOAD	40
Figura 16- Inserção do maior valor do identificador principal na Tabela DW_LOAD	40
Figura 17- Modelo da Tabela DIM_TEMPO	41
Figura 18- Modelo de Visão Materializada VM_KPI1	42
Figura 19- Resultado da evolução de tempo para a carga inicial Erro! Indicador não definido.	
Figura 20- Resultado da evolução de espaço alocado em disco.....	45
Figura 21- Analogia entre o tempo de resposta V_KPI1 X VM_KPI1	46

LISTA DE TABELAS

Tabela 1- Dados, Informações e Indicadores.....	21
Tabela 2- Comparação das características de dados operacionais e de suporte a decisões	26
Tabela 3- Estrutura de dados de indicadores	35
Tabela 4- Carga inicial e tempo médio de execução	44

LISTA DE ABREVIATURAS E SIGLAS

DW	Data Warehouse
ERP	Enterprise Resource Planning
ETL	Extract, Transform and Load
ISO	International Organization for Standardization
KPI	Key Performance Indicator
OLAP	On-Line Analytical Processing
OLTP	On-Line Transaction Processing
SGBD	Sistema de Gerenciamento de Banco de Dados

Capítulo 1 - INTRODUÇÃO

Neste capítulo, introduz-se o contexto onde está inserida esta dissertação. Para tal, apresentam-se a motivação, o problema de pesquisa, os objetivos a serem alcançados, bem como a organização dos demais capítulos do trabalho.

1.1 Motivação

No mundo empresarial em constante evolução, a competição entre as empresas se intensificou. Neste cenário de grandes transformações, é fundamental que as instituições possam armazenar grande volume de dados, provenientes e/ou extraídos dos mais diversos sistemas transacionais. Assim, uma solução para a correta organização e manipulação dos dados de decisões é o uso de um *Data Warehouse*. Os *Data Warehouses* são armazenados em estruturas otimizadas para a velocidade de análise e consulta, eles são projetados com o objetivo de suportar extração e processamento eficiente de dados.

Como a competição se intensificou e a informação tornou-se facilmente disponível, as medidas financeiras de desempenho têm sido vistas como medidas de desempenho insuficiente e as medidas não financeiras tornaram-se cada vez mais valiosas como novas e volumosas fontes de informação, pois facilitam aos gestores a tomada de decisões estratégicas (Finger e Mussi, 2014). Além disso, ajudam a prever o desempenho em longo prazo melhorando assim o planejamento e o controle (Lingle e Schiemann, 1999).

Desta forma, para manipular essa base de dados volumosa faz-se necessário intercalar os dados para obtenção dos indicadores chaves de desempenho (em inglês, *Key Performance Indicator* – KPI). Esses indicadores são os resultados de uma medida ou várias medidas que tornam possível a assimilação da evolução do que se aspira avaliar a partir de suas particularidades (Moreira, 2002; Rob e Coronel, 2011), norteando, assim, as tomadas de decisões, pois mensuram de forma clara os resultados obtidos de acordo com as necessidades de informação colocadas pela empresa.

Apesar das tecnologias e inovações no mercado observa-se ainda uma prática comum em pequenas e médias empresas a obtenção dos indicadores de negócios de forma manual. Este processo ocorre através da emissão de relatórios do sistema corporativo, que logo após os dados serem obtidos são digitados em

planilhas eletrônicas, para serem reprocessados e gerados os gráficos de acordo com os seus indicadores. Em seguida, os gestores analisam os resultados para as devidas apresentações em reuniões mensais, bimestrais e/ou semestrais. Ao término desse processo são realizadas as tomadas de decisões e projeções futuras para a gestão de negócios.

Para obtenção automática desses indicadores pode-se utilizar processo de extração, transformação e carregamento de dados (em inglês *Extract Transform Load* - ETL), uma vez que, tal processo, é responsável por coletar, filtrar, integrar e agregar dados transacionais a serem salvos em um armazenamento de dados otimizados para o suporte a decisões (Rob e Coronel, 2011).

Para manipulação de dados transacionais, tem-se as aplicações de processamento de transações on-line (em inglês *On-Line Transaction Processing* - OLTP), que são sistemas projetados para suportar processamento de dados gerados no cotidiano, por exemplo, sistemas corporativos (Controle de Estoque, Compras e Faturamento, Financeiro, Educacional, Folha de Pagamento dentre outros). Têm-se, também, as aplicações de processamento de transações analíticas on-line (em inglês, *On-Line Analytical Transaction* - OLAP), que são sistemas utilizados para analisar grandes volumes de dados nas mais diversas perspectivas com objetivo de auxiliar as tomada de decisões.

Porém, para melhorar o tempo de processamento de consultas em sistemas OLAP, utiliza-se uma técnica denominada de visão materializada. Dessa forma, as consultas são executadas de forma mais rápida, reduzindo o tempo de processamento (Gupta e Mumick, 1995).

Entretanto, a automatização dos indicadores de negócios é um desafio de gestão para as pequenas e médias empresas. Mediante esta inquietação, o presente trabalho propõe apresentar uma abordagem capaz de minimizar o tempo, mitigar erros e reduzir o espaço alocado em disco para obtenção dos indicadores de negócios. A seguir especifica-se os objetivos deste trabalho e apresenta-se os capítulos que compõem a dissertação.

1.2 Objetivos

Propor uma abordagem para que o processo de ETL seja integrado aos indicadores de negócios, através de visões materializadas em aplicações de *Data Warehouse* para auxiliar na tomada de decisão.

Neste contexto, delimita-se os seguintes objetivos específicos propostos:

- Comparar o processo de carga inicial da abordagem atual com a abordagem proposta, através do processo de ETL integrado aos indicadores de negócios;
- Realizar testes experimentais usando base de dados relacional real compartilhada por instituição de ensino básica do Brasil;
- Analisar se a utilização de indicadores de negócios trouxe vantagens, tais como: melhoria no tempo de acesso para obtenção dos indicadores, mitigar erros durante o processo e redução do espaço alocado em disco.

1.3 Formulação do Problema

A área de pesquisa deste trabalho é integração do processo de ETL com os indicadores de negócios através de visões materializadas em aplicações de *Data Warehouse* para auxiliar na tomada de decisão. A título de ilustração, apresenta-se um exemplo de um cenário típico do problema de obtenção de forma manual dos indicadores de negócios em pequenas e médias empresas. A **Figura 1** apresenta um esquema deste cenário.

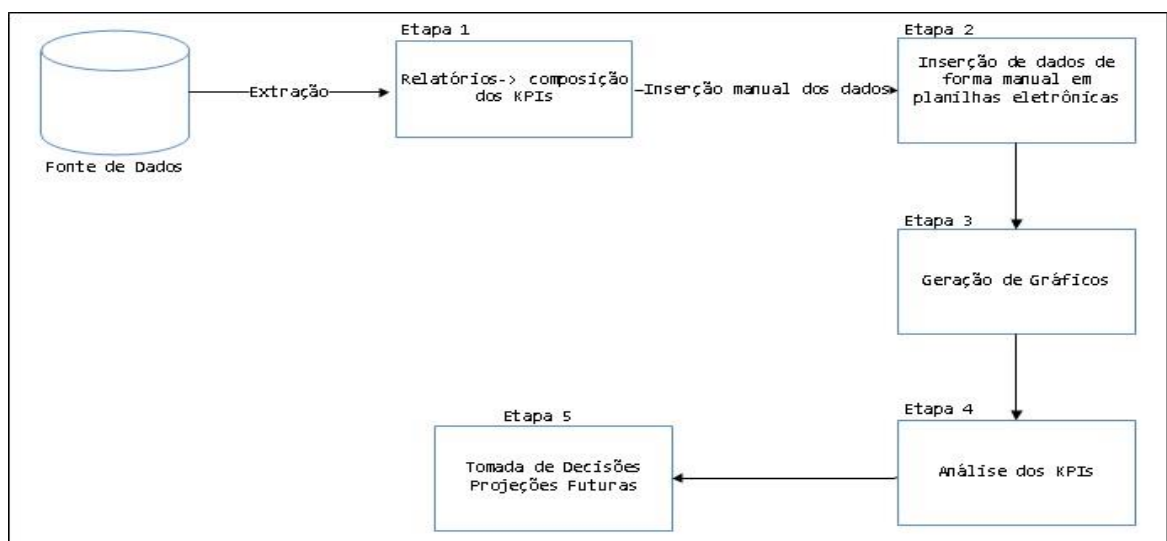


Figura 1 Cenário obtenção dos indicadores de negócios

De acordo com este cenário (**Figura 1**), pode-se verificar como ocorre o processo de obtenção dos indicadores de negócios. A seguir são descritas as etapas:

- 1) Para o processo de elaboração/disponibilização dos indicadores de negócios prepara-se a planilha eletrônica para receber os dados desse processo, conforme informações da empresa estima-se o tempo médio de aproximadamente 02 (duas) horas por indicador.
- 2) São gerados relatórios da base de dados que compõem os KPIs;
- 3) Logo após são digitados os dados em planilhas eletrônicas para atualização, esse processo é realizado no tempo mínimo de 10 (dez) minutos;
- 4) Após a digitação dos dados, são gerados gráficos e em seguida os gestores analisam os dados para as devidas apresentações;
- 5) Para a finalização do processo, são realizadas reuniões onde são tomadas as decisões e projeções futuras para a gestão de negócios.

Pode-se notar alguns problemas para obtenção desses indicadores, conforme apresentado na **Figura 1**.

- 1) Dado a manipulação dos dados diretamente pelos usuários, há possibilidade de erro é inevitável, tornando o processo impreciso em suas informações. Consequentemente as tomadas de decisões, as projeções para a gestão de negócios ficam comprometidas devido esta forma manual de obtenção dos KPIs;
- 2) Alta demanda de tempo de alocação dos colaboradores envolvidos no processo, pois como observado, é realizado em diversas etapas (coleta/extração, digitação/geração, apresentação)
- 3) Aumento da alocação de espaço em disco, devido ao processo de disponibilização dos indicadores.

Este cenário corrobora a necessidade de uma abordagem para que os indicadores de negócios sejam obtidos de forma automática através do processo de ETL integrado aos indicadores de negócios através de visões materializadas em aplicações de *Data Warehouse* para auxiliar na tomada de decisão.

1.4 Organização da Dissertação

Os estudos e análises realizados no desenvolvimento deste trabalho se organiza com a seguinte estrutura: No capítulo inicial, contempla a parte introdutória do trabalho, na qual se abordam a motivação, objetivos da pesquisa, apresentação do problema e a estrutura organizacional da dissertação.

Em seguida, o capítulo 2 apresenta a fundamentação teórica, relacionada ao levantamento de literatura atualizada sobre as temáticas envolvidas, para dar embasamento teórico na discussão dos conceitos e tecnologias utilizados no contexto do trabalho. Este levantamento será feito através de livros e artigos publicados em revistas científicas especializadas. Neste capítulo ainda será discutido os trabalhos correlacionados através de pesquisa buscando informações sobre técnicas de integração de processo ETL com a tecnologia OLAP, visões materializadas e uso de indicadores chaves de desempenho. Este aporte foi fundamental para a elaboração da proposta de uma nova abordagem sobre este processo.

O capítulo 3 descreve os procedimentos metodológicos, onde é apresentada a arquitetura do modelo proposto nesta pesquisa e feita sua descrição. São descritas as etapas do pré-processamento, transformação dos dados, implementação de indicadores de negócios e elaboração das visões materializadas.

O capítulo 4 apresenta os resultados obtidos em que são analisados a evolução do tempo de inserção da carga inicial, espaço alocado em disco e o tempo de resposta das visões materializadas.

O capítulo 5 descreve a conclusão, que ampliam as discussões referentes a esta pesquisa, apresenta novos desafios para trabalhos futuros que merecem ser discutidos, avaliados e pesquisados colaborando com a área de pesquisa.

Capítulo 2 – FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta o aporte teórico utilizado como fundamento deste trabalho. Os conceitos apresentados a seguir nortearam as atividades dessa pesquisa, bem como influenciaram a concepção da abordagem.

2.1 Indicadores Chaves de Desempenho

A frase de Deming (1990), estatístico, norte-americano, “não se gerencia o que não se mede” ou “medir é importante: o que não é medido não é gerenciado” (Kaplan, Norton, 197). Aponta o objetivo e o foco de um indicador de desempenho, mensura o que está sendo executado, ajuda na gestão de tomadas de decisões, na análise crítica do desempenho da organização e para o planejamento.

Os Indicadores Chaves de Desempenho (do inglês *Key Performance Indicator* - KPI), representam um conjunto de medidas com foco sob os aspectos de desempenho organizacional que são os mais críticos para o sucesso atual e futuro da organização (Parmenter, 2007, p. 3). Segundo Rob e Coronel (2011), o monitoramento dos negócios é primordial para a saúde da empresa, seja para saber onde está atualmente e para onde pretende chegar. Estes indicadores são medidas quantificáveis (numéricas ou baseadas em escala) que avaliam a eficiência ou o sucesso da empresa em alcançar seus objetivos.

Segundo FNQ (2015, p. 4), o KPI é uma informação quantitativa ou qualitativa que expressa o desempenho de um processo, em termos de eficiência, eficácia ou nível de satisfação e que, em geral, permite acompanhar sua evolução ao longo do tempo e compará-lo com outras organizações.

Os KPIs são de grande valia, pois mensuram o sucesso ou o fracasso de uma empresa. Sem o auxílio deles torna-se difícil saber se o negócio está indo realmente em direção às metas previamente definidas. Eles não dizem o que fazer para melhorar, apenas reflete a situação atual. O mais importante, neste caso, não é a medição e, sim, as ações que serão tomadas a partir desses indicadores, ou seja, o gestor deve tornar seu processo de decisão menos aleatório e mais científico (Harrington, 1993).

Além de promover a cultura para a excelência, medir os resultados sistematicamente de forma estruturada é fundamental para a gestão de uma organização, uma vez que possibilita a comparação. Se as organizações quiserem

sobreviver e prosperar na era do conhecimento e da informação, deverão utilizar sistemas de gestão e medição derivados de suas estratégias e capacidades.

A comparação, essa por sua vez está relacionada: ao passado (série histórica), ao referencial de desempenho, ao compromisso assumido e a meta de desempenho. A possibilidade de mensurar o desempenho por meio de indicadores é uma necessidade. Vários fatores contribuem para isso, conforme FNQ (2015, p. 3):

- Grau de exigência das partes interessadas – *stakeholders* - que aumenta a necessidade de um processo de medição, objetivo, sistemático e transparente, o qual não seja restrito a indicadores econômico-financeiros;
- A prática de remuneração variável com base em um sistema estruturado de indicadores organizacionais está cada vez mais disseminada, gerando a necessidade de critérios corporativos únicos de medição;
- Aumento da velocidade e da qualidade na tomada de decisões e, conseqüentemente, nos seus efeitos. A organização precisa ser cautelosa ao avaliar o impacto sistêmico delas no desempenho global.

Segundo Parmenter (2007), a partir da análise extensa e das discussões com mais de 1.500 participantes em oficinas de KPI, cobrindo tipos de organização nos setores público e privado, apresenta 07 (sete) características dos KPIs:

- Refletir objetivos da empresa como um todo;
- Devem ser utilizados pela direção para administrar e tomar decisões, pois tem caráter estratégico;
- Ter alto grau de aderência ao negócio da empresa;
- Ter relevância em todos os níveis da empresa;
- Baseados em dados confiáveis e mensuráveis;
- Fácil de entender (ou pelo menos rápido de ser explicado);
- Servir como insumo para uma ação ou um plano de ações.

Há uma diferença nos conceitos dados, informações e indicadores (**Tabela 1**). Um dos motivos dessa diferenciação é permitir comparações inerentes a outras organizações. Indicadores genuínos já carregam em sua essência uma possibilidade comparativa convincente, segundo afirma Peter Drucker (2002): “Mas cada vez mais os desafios não são técnicos, ao contrário, eles estarão ligados a conversão de dados em informações utilizáveis”. A maioria das métricas adotadas nas estruturas

de indicadores das empresas tem como base as informações, sendo os indicadores um bem mais escasso do que se imagina.

Tabela 1- Dados, Informações e Indicadores

Dados	Informações	Indicadores
Disponíveis para manipulação no banco de dados	Organizadas e já manipuladas em primeiro nível	Manipulados matematicamente por meio de fórmulas (divisão)
Abundantes e armazenados em sua totalidade	Selecionadas em formatos de telas e/ou relatórios	Parametrizados em formatos de gráficos lineares
Viabilizados por meio de coleta de dados	Viabilizadas por meio de <i>softwares</i> gerenciais	Viabilizados por meio de contatem
Não têm foco na gestão	Com foco abrangente e dispersivo	Com foco no que é relevante

Fonte: FNQ, 2015

A seguir, alguns exemplos de KPIs a partir de Rob e Coronel (2011):

Gerais: Medidas anuais de lucro por linha de negócios, vendas comparáveis de lojas, rotatividade e *recalls* de produtos, vendas por promoção, por funcionário dentre outros.

Financeiros: Ganhos por ação, margem de lucro, receita por funcionários, porcentagem de vendas em relação a contas a receber, ativos à venda, lucratividade, inadimplência, rentabilidade dentre outros.

Recursos humanos: Candidatos a cargos abertos, absenteísmo, rotatividade e longevidade de funcionários dentre outros.

Educação: Taxa de formandos, números de calouros ingressantes, taxa de retenção de alunos, dentre outros.

2.2 Processo ETL

Segundo Inmon (1997), o processo de ETL às vezes é visto como uma simples extração de dados do ambiente operacional para a carga no ambiente do *Data Warehouse*. Mas, o que parece simples na verdade é a fase mais importante e a mais complexa na implantação deste ambiente, que ocupa cerca de aproximadamente 60% a 70%. Selecionar os dados relevantes de várias bases

heterogêneas e trazê-los para o padrão definido no *Data Warehouse* é uma tarefa onerosa. Os processos de Extração, Transformação e Carga (ETL, do inglês *Extract Transform Load*), são fundamentais para a preparação dos dados que serão armazenados no *Data Warehouse* (Erro! Fonte de referência não encontrada.).



Figura 2- Processo de ETL, (Brave River Solutions)

Para Primak (2008), as ferramentas de ETL possibilitam cinco operações principais. A primeira delas refere-se a extração dos dados que podem estar em fontes internas (sistemas transacionais, banco de dados etc) ou externas (em sistemas fora da empresa). Em seguida, é preciso fazer a limpeza e transformação dos dados. Nessa etapa são corrigidas algumas imperfeições contidas na base de dados transacional com o objetivo de fornecer dados precisos, não redundantes e com qualidade, permitindo uma análise fundamentada nos dados mais próximos do real.

A terceira operação destina-se ao processo de transformação do dado. Geralmente, os dados vêm de vários sistemas diferentes e por isso se faz necessário padronizá-los. A quarta etapa diz respeito ao processo de carga do *Data Warehouse* que em geral é feita partir de um banco de dados temporário, no qual os dados armazenados já passaram pela limpeza e integração. E, finalmente, a etapa da atualização dos dados (*Refresh*). Sabe-se que a todo momento ocorrem alterações na base de dados transacional e devem ser passadas para o *Data Warehouse*.

Primak (2008), também sugere fases necessárias para a implantação do processo de ETL (**Figura 3**):

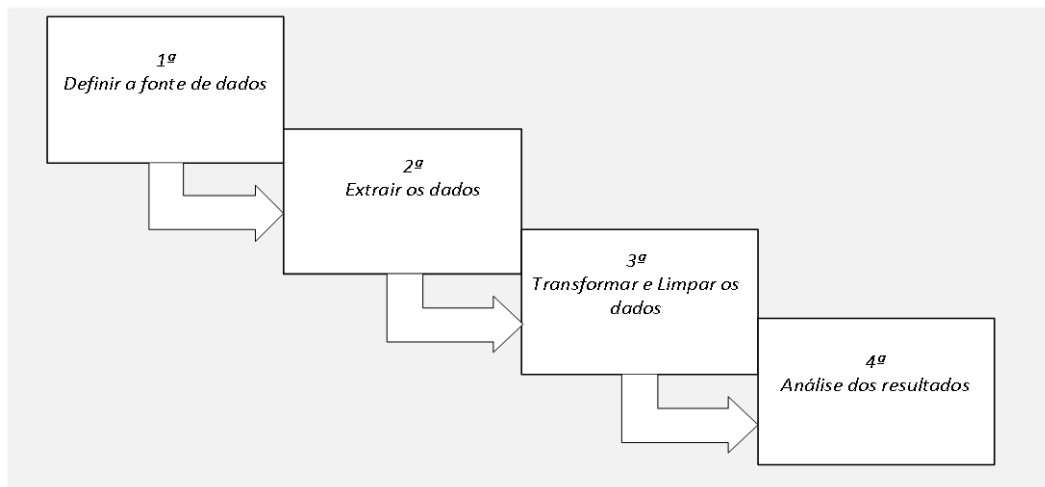


Figura 3- Fases de implantação do processo de ETL (Primak, 2008)

Nesta fase, quando obtêm-se os dados de uma fonte, os mesmos podem possuir inconsistências, por exemplo, durante a venda de produtos para cliente, a principal preocupação do vendedor é realizar a venda, se o cliente não tiver o CPF em mãos ele cadastra um número qualquer, deste que o sistema aceite, um dos mais utilizados 77777777-77 devido a emissão de notas fiscais. Numa hipótese o diretor financeiro desta empresa consulta o seu *Data Warehouse* para verificar quais são seus maiores clientes e aparece em primeiro lugar o cliente que tem o CPF 77777777-77? Por isso, nessa fase, verifica-se a compatibilidade e consistência entre os dados.

Na fase seguinte, temos de fazer a limpeza e transformação de dados, pois os dados provêm de vários sistemas, podendo ocasionar, formatos diferentes para o mesmo dado. Por exemplo, informações sobre o sexo do cliente pode estar armazenada nos seguinte formato: “M” para Masculino e “F” para Feminino, porém em algum outro sistema está armazenado “H” para Masculino e “M” para Feminino e assim sucessivamente. Então, quando realizado o processo de ETL, os dados serão transformados e padronizados, isso facilitará a análise dos mesmos a serem recuperados.

Com o volume de dados é muito grande, há muitos casos que não temos condições de processar as extrações e transformações na janela de tempo em que o *Data Warehouse* não está sendo usado, então deve-se que fazer uso do que se denomina de *staging area* (é um repositório de dados que serve para integrar vários tipos de dados em um único formato), para que se consiga executar os processos com sucesso.

A seguir são apresentados alguns dos fatores que devem ser analisados antes de começar a fase de extração de dados (Primak, 2008):

- A extração de dados do ambiente operacional para o ambiente de *Data Warehouse* demanda uma mudança na tecnologia. Os dados são transferidos de várias fontes de dados para *Data Warehouse*;
- A seleção dos dados do ambiente operacional pode ser muito complexa, pois muitas vezes é necessário selecionar vários campos de um sistema transacional para compor um único campo no *Data Warehouse*;
- Outro fator importante é que dificilmente há o modelo de dados dos sistemas antigos, e, se existem, não estão documentados;
- Os dados são reformatados. Por exemplo, um campo data do sistema operacional do tipo dd/mm/aaaa pode ser passado para o outro sistema do tipo ano e mês como aaaa/mm;
- Quando há vários arquivos de entrada, a escolha das chaves deverá ser feita antes que os arquivos sejam intercalados. Isso significa que, se várias chaves são usadas nos diferentes arquivos de entrada, então, opta-se por apenas uma dessas estruturas;
- Os arquivos devem ser gerados obedecendo a mesma ordem das colunas no ambiente de *Data Warehouse*;
- Pode haver diversos resultados. Dados podem ser produzidos em diferentes níveis de resumo pelo mesmo programa de geração das cargas;
- Valores *default* devem ser fornecidos. Isso significa, quando o campo no *Data Warehouse* que não possui fonte de dados, então a solução é definir um valor padrão para este campo.

Dessa forma, para garantir uma qualidade dos dados, é importante observar suas características mais relevantes são: **unicidade**, evitando assim duplicações de dados; **precisão**, os dados não podem perder suas características originais assim que são carregados para o *Data Warehouse*; **completude**, não gerando dados parciais de todo o conjunto relevante às análises; e **consistência**, ou seja, os fatos devem apresentar consistência com as dimensões que o compõem. É necessário que os dados fiquem em uma forma homogênea para serem carregados no *Data Warehouse* (Kimball, 1998).

Observa-se as seguintes questões que devem ser consideradas em relação à carga de dados (Goodyear, 1999):

A janela de tempo disponível para a realização da carga deve ser planejada, analisada e determinada, pois, sendo um ambiente de consultas, não há necessidades que o *Data Warehouse* permaneça operacional 24 (vinte e quatro) horas por dia. Mas, muitos projetistas - motivados mais pelo estímulo de auto superação do que propriamente por necessidade de negócio, estão definindo arquiteturas capazes de diminuir (e muito) a janela *off-line* do sistema, podendo até mesmo manter o *Data Warehouse* operacional durante todo o dia, sem prejuízo para a carga de dados, assim sendo, é importante ressaltar que esta janela depende muito da área de negócios da empresa.

A frequência com que as operações de cargas serão realizadas também deve ser analisada e determinada. Isto é diretamente proporcional à taxa de crescimento ou alteração das tabelas de dados do ambiente operacional. Pode-se citar, como exemplo, que algumas tabelas devem ser carregadas somente uma vez por mês, enquanto que outras que sofrem mudanças mais rapidamente precisam ser carregadas diariamente.

O desempenho da carga é altamente dependente da disponibilidade de recursos de sistemas. Isto implica em limitações de memória e disco, que são as causas mais comuns de degradação de desempenho e falha na carga.

Apesar de existirem várias ferramentas de ETL no mercado sejam *open-source* (*Apator*, *Pentaho Community*, *SpagoBI*, *KETL*, *Palo* dentre outras) ou proprietárias (*Oracle Business Intelligence Standard*, *Microsoft Business Intelligence*, *SAP Business Intelligence* e *IBM Cognos*), ainda tem-se a necessidade de criar rotinas de carga de dados para atender determinadas situações que poderão ocorrer.

2.3 Data Warehouse

Para entender o que é um *Data Warehouse*, é importante salientar uma analogia com o conceito tradicional de banco de dados.

Date (2003), o banco dados, pode ser considerado como o equivalente eletrônico de um armário de arquivamento; ou seja, é um repositório ou recipiente para uma coleção de arquivos de dados computadorizados. Os usuários podem solicitar que o sistema realize diversas operações envolvendo tais arquivos, por exemplo:

- Acrescentar novos arquivos ao banco de dados;

- Inserir dados em arquivos existentes;
- Buscar dados de arquivos existentes;
- Excluir dados de arquivos existentes;
- Alterar dados em arquivos existentes;
- Remover arquivos existentes em banco de dados.

Em consideração a esta definição sobre banco de dados, pode-se dizer que um *Data Warehouse*, conforme, Inmon (1997), é uma coleção de dados, integrado, orientados por assunto, variável no tempo e não volátil, que tem por objetivo dar suporte aos processos de tomada de decisões.

Para Primak (2008), os bancos de dados armazenam dados operacionais necessários para as operações diárias da empresa, são utilizados por todos os funcionários para registrar e executar operações pré-definidas, por isso seus dados podem sofrer constantes mudanças, conforme as necessidades da empresa.

Um *Data Warehouse* armazena dados analíticos, destinados as necessidades da administração no processo de tomada de decisões. Elmasri, Navathe (2005), eles têm a característica distinta de que são direcionados principalmente para aplicação de apoio as decisões. Um *Data Warehouse* armazena informações históricas e conseqüentemente uma grande capacidade de processamento e armazenamento dos dados.

Conforme, Rob e Coronel (2011), os dados de suporte a decisão diferem dos operacionais em três áreas operacionais: período de tempo, granularidade e dimensionalidade (**Tabela 2**).

Tabela 2- Comparação das características de dados operacionais e de suporte a decisões

Característica	Dados Operacionais	Dados de Suporte a Decisão
Atualidade dos dados	Operações correntes Dados em tempo real	Dados históricos Retratos instantâneo dos dados da empresa Componente temporal (semana/mês / ano)
Granularidade	Dados detalhados no nível indivisível	Dados resumidos
Nível de resumo	Baixo; alguns resultados agregados	Alto; vários níveis de agregação
Modelo de dados	Altamente normalizado Principalmente SGBD relacional	Não normalizado Estrutura completa Principalmente em SGBD

		multidimensional, embora com alguns aspectos relacionais
Tipo de transação	Principalmente atualizações	Principalmente consulta
Volumes de transação	Alto volume de atualizações	Cargas periódicas e cálculos resumidos
Velocidade de transação	As atualizações são fundamentais	As recuperações são fundamentais
Atividade de consulta	De baixa a média	Alta
Escopo de consulta	Pequena amplitude	Grande amplitude
Complexidade da consulta	De simples a média	Muito complexa
Volumes de dados	Centenas de megabytes, até gigabytes	Centenas de gigabytes, até terabytes

Fonte: Rob e Coronel (2011)

O *Data Warehouse*, consiste em organizar os dados corporativos de maneira integrada, orientada por assunto, histórica, variável com o tempo e gerando uma única fonte de dados. Para compreender melhor essa organização, pode-se verificar mais detalhadamente em seus componentes (Inmon, 1997):

- Integrado: O *Data Warehouse* é um banco de dados consolidado e centralizado que integra dados provenientes de toda a organização e de várias fontes, com diversos formatos. Integração significa que todas as entidades comerciais, elementos e características de dados e métricas de negócios estão descritas do mesmo modo em toda a empresa;
- Orientado por assunto: Os dados em *Data Warehouse* são dispostos e otimizados de modo a fornecerem respostas a perguntas provenientes de diversas áreas funcionais de uma empresa. São organizados por temas, como vendas, marketing, finanças, distribuição e transporte. Para cada tema o *Data Warehouse* contém assuntos de interesse específico – produtos, clientes, departamentos, regiões, promoções, dentre outros;
- Variável no tempo: Ao contrário dos dados operacionais que focam nas transações correntes, os dados em *Data Warehouse* representam o fluxo através do tempo. Também é variável no tempo, no sentido de que, uma vez que os dados são carregados periodicamente no *Data Warehouse*, todas as agregações dependentes do tempo são recalculadas;
- Não volátil: Uma vez inserido um dado no *Data Warehouse*, ele não será removido. Como os dados no *Data Warehouse* representam o histórico da empresa, os dados operacionais, que constituem o histórico de curto

prazo, são sempre adicionados a ele. Como os dados nunca são excluídos e novos dados são inseridos continuamente, o *Data Warehouse* está sempre crescendo. Por isso, o SGBD (Sistemas de Gerenciamento de Banco de Dados) deve ser capaz de dar suporte a banco de dados com vários gigabytes ou, até mesmo, *terabytes*, operando em *hardware* com diversos processadores.

Para viabilizar um projeto de *Data Warehouse*, faz-se necessário a presença de uma série de componentes. Turban *et al* (2009) desenvolveram os seguintes componentes e suas interações (**Figura 4**).

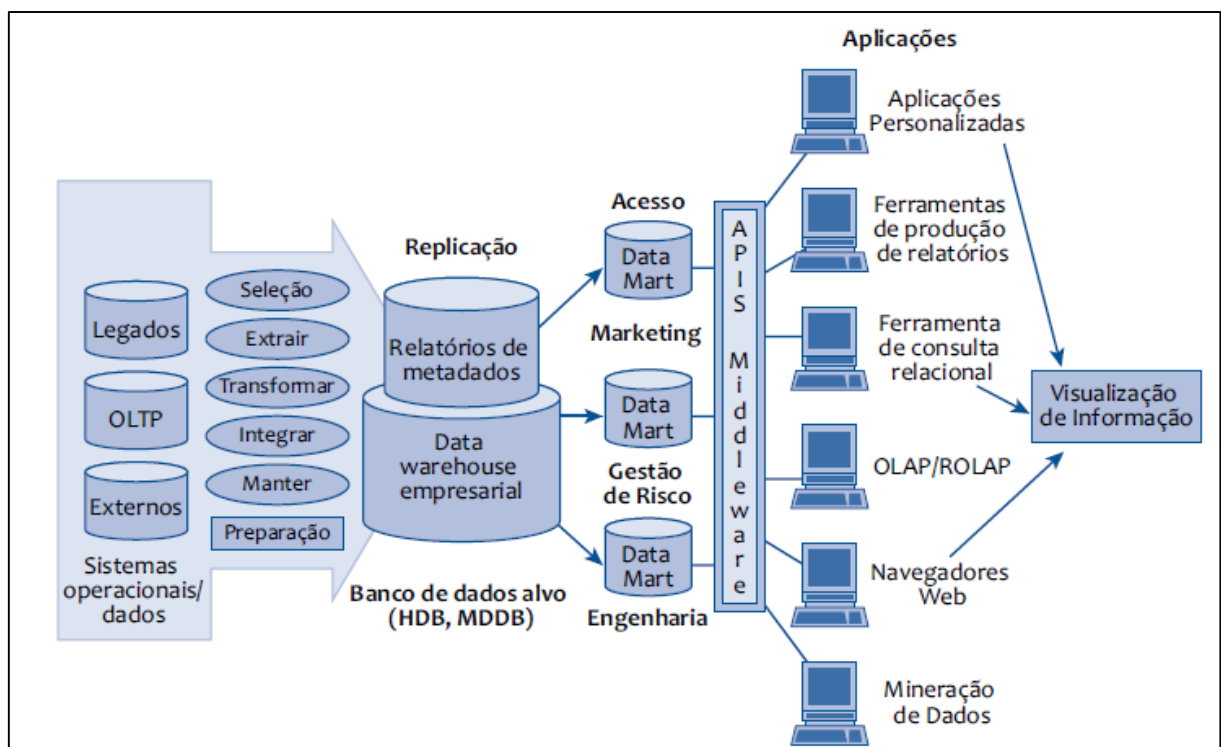


Figura 4- Principais componentes (Turban et al, 2009).

A partir da análise da **Figura 4** são apresentados mais detalhes sobre a descrição dos componentes:

- **Fontes de dados:** é composta por várias bases de dados transacionais e espalhadas pela organização, que fazem parte do sistemas de informações operacionais, tais como: Sistemas legados, Sistemas ERP (Sistemas integrados de gestão);
- **Extração de dados:** É denominado processo de ETL (**seção 2.2**);
- **Staging area:** é uma camada intermediária entre os dados operacionais e a carga do *Data Warehouse*, tem a função de limpeza dos dados antes que eles sejam transportados para o *Data Warehouse*;

- **Metadados:** Facilitam a recuperação e organização dos dados armazenados;
- **Ferramentas de *Middleware*:** Interface de acesso, para que outros serviços possam interagir com as informações armazenadas.

Portanto, o *Data Warehouse* normalmente é um banco de dados apenas de leitura, otimizado para processamento de análises e consultas. A **Figura 5** ilustra a arquitetura do *Data Warehouse* criado a partir dos dados contidos em sistemas aplicativos.

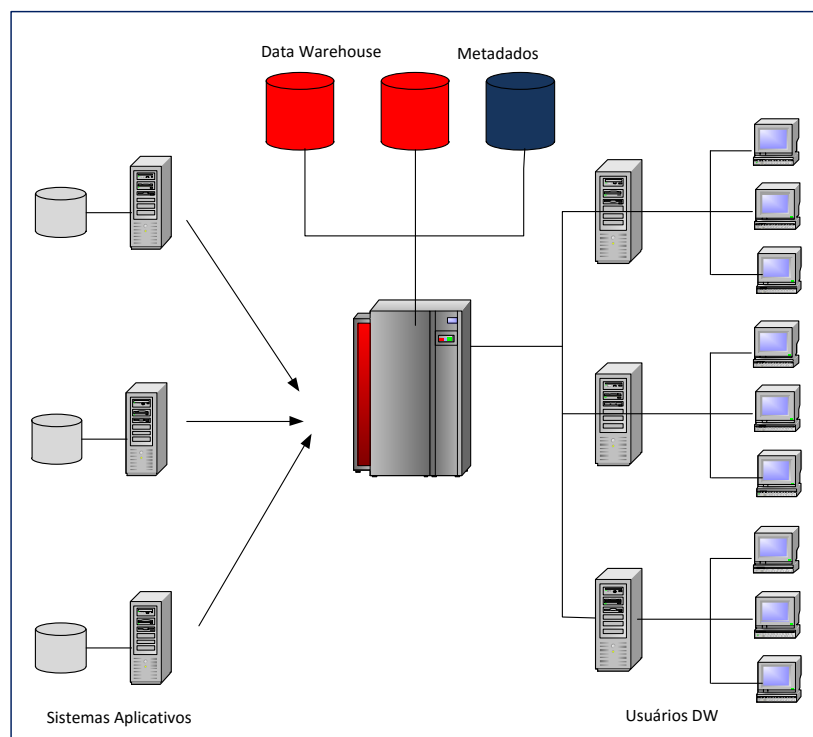


Figura 5- Arquitetura do Data Warehouse (Primak, 2008)

É comum o uso de cargas iniciais em *Data Warehouse*, transportando todo o conteúdo das tabelas que farão composição do seu ambiente. Esta é a última parte deste processo de povoamento. Neste momento serão executadas as últimas etapas de preparação dos dados, a carga propriamente dita e algumas atividades posteriores à carga, como o tratamento dos dados rejeitados e o processo de certificação da qualidade dos dados carregados.

Basicamente, existem dois tipos de gerações de carga:

- **Bulk Load:** carrega o conjunto inteiro de registros (OLTP) no *Data Warehouse*. Em geral, os banco de dados oferecem recursos

apropriados que são utilizados para realização de forma eficiente estas operações de inserção de dados em massa a partir da linha de comando;

- **SQL Load:** identifica as alterações ocorridas desde a última carga (OLTP) e as carrega para dentro do Data Warehouse.

No estudo de caso deste trabalho, a abordagem atual utiliza *bulk load* e *sql load* para carregamento dos registros, enquanto a abordagem proposta utilizou *sql load* baseadas em indicadores de negócios.

2.4 Visão Materializada

Na literatura, o conceito de visão, é explicado por Ramakrishnan (2008) como uma consulta, cuja Tabela resultado não é armazenada explicitamente no banco de dados. As tuplas são calculadas conforme requisitado pelo sistema. Portanto, uma visão é uma definição de consulta sobre um conjunto de Tabelas do banco de dados que é recalculada cada vez que a visão é referenciada.

Para Rob e Coronel (2011), a visualização relacional possui várias características especiais, tais como:

- É possível utilizar o nome de uma visualização em qualquer posição de comandos de SQL em que deva ser inserido o nome da Tabela;
- São atualizadas dinamicamente, ou seja, a visualização é recriada a cada vez que for solicitada;
- Fornecem um nível de segurança no banco de dados, pois podem restringir o acesso dos usuários apenas com as colunas e linhas especificadas;
- Podem ser utilizadas como base para relatórios.

Em analogia com o conceito de visão, à visão materializada, os resultados da consulta são armazenados no banco de dados como tabelas concretas que podem ser atualizadas a partir das Tabelas originais. Dessa forma, já que uma visão materializada é armazenada como uma Tabela real qualquer operação que é permitida em uma Tabela normal também pode ser aplicada a uma visão materializada.

De acordo com Kimball (2002) e Silberschatz (2006), as visões materializadas possuem várias características, tais como:

- Melhoram o processamento de consultas;
- Sumarização dos dados;

- Controle de desempenho em aplicações de *Data Warehouse*, aperfeiçoando o tempo gasto em acessos para escrita, pois o resultado da execução já está armazenado e podem ser construídos índices sobre o resultado materializado.

O uso de visões materializadas pode ser mais vantajoso do que índices em consultas com agregações e agrupamentos, pois elas já armazenam o resultado das agregações não precisando ser calculada quando for referenciada (Carvalho, 2011). Para compreender essa diferença, será analisado o seguinte exemplo. Observa-se a consulta VC e a visão VMC sobre a Tabela Movimento, conforme **Figura 6**:

Consulta VC	Visao VMC
<pre>Select idmov From movimento Group By idmov Having Sum(valortotal)>5000</pre>	<pre>Select idmov, Sum(valortotal) From movimento Group By idmov</pre>

Figura 6- Consulta VC e Visão VMC

Considerando a existência de um índice sobre o campo *idmov* da Tabela Movimento, ele seria usado para agrupar a consulta VC, porém ainda precisaria ser feito a agregação “sum” do *having*. Considerando a existência da visão VMC, seria necessário apenas percorrer a visão materializada VMC para obter o mesmo resultado da consulta VC, sem ser necessário fazer agrupamento e agregação.

Entretanto, há desvantagem das visões materializadas (Quass, 1997; Silberchatz, 2006) elas precisam de atualização de dados, quando existem mudanças nos dados das tabelas referenciadas na consulta que define uma visão materializada, esta precisa ser renovadas para refletir essas mudanças.

É importante ressaltar que os sistemas de bancos de dados modernos oferecem suporte direto para manutenção incremental de visão. Ou seja, é realizada assim que ocorrer uma atualização, como parte da transação de atualização. Alguns SGBDs também admitem a manutenção de visão adiada, dessa forma, é atualizada em outra ocasião, por exemplo, as atualizações podem ser coletadas durante o dia, e as visões materializadas podem ser atualizadas a noite (Silberchatz, 2006).

Assim, o uso de visões materializadas nos *Data Warehouses* tem como objetivo minimizar o tempo de resposta destas consultas e os custos de manutenção das visões. Na abordagem proposta todas as visões materializadas foram implementadas através dos indicadores de negócios.

2.5 Trabalhos Relacionados

Para embasamento do presente trabalho, após a realização de um levantamento sobre trabalhos de objetivos similares, foram verificados alguns trabalhos com características complementares, e que serviram de suporte para o desenvolvimento desta abordagem. A pesquisa procurou obter informações sobre técnicas de integração de processo ETL, novas tecnologias, visões materializadas e uso de indicadores chaves de desempenho em *Data Warehouse*.

Um primeiro trabalho que chama bastante atenção é o Quass (1997), descreve uma pesquisa em que se propõe a uma abordagem de visões materializadas para integração do processo de ETL em aplicações Data Warehouse. Esta abordagem pode prover uma rápida e eficiente propagação e alterações de sistemas de origem OLTP para os sistemas de Data Warehouse OLAP. O uso de visões materializadas por reescrita de consulta para atualização de informações e estão disponíveis para suporte quase em tempo real para análise/mineração e consequentemente benefícios nas tomadas de decisões.

A abordagem na forma descrita no trabalho tem características interessantes e desejáveis ao desenvolvimento desta proposta, principalmente no que tange à sua arquitetura. Contudo, o mesmo não contempla o uso de indicadores de negócios no processo de ETL e na concepção das visões materializadas em aplicações de *Data Warehouse*.

De forma equivalente, Italiano (2000), refere o uso de visões materializadas nos *Data Warehouses* tem como objetivo minimizar o tempo de resposta destas consultas e os custos de manutenção das visões. Destacam-se alguns aspectos importantes como a seleção das visões mais adequadas, dado um conjunto de pesquisas comuns e as características dos processos de manutenção das visões de forma dinâmica e incremental.

Outro trabalho interessante neste campo de pesquisa Neves(2009), que apresenta uma abordagem que automatiza a seleção de visões a materializar, para *Data Warehouses*, levando em consideração parâmetros do banco de dados transacional e as consultas dos usuários, identificadas na fase de definição do projeto conceitual do *Data Warehouse*, em consequência dos requisitos de análise dos usuários.

Considerando os trabalhos pesquisados, é interessante ressaltar que apesar de constituírem iniciativas interessantes, nenhum se presta completamente a tarefa de fazer com que o processo de ETL seja integrado aos indicadores de negócios através de visões materializadas, portanto propõe-se nesta dissertação a investigar tecnologias que possam contribuir na tarefa dessa integração para a obtenção de forma automática dos indicadores de negócios para auxiliar na tomada de decisão.

Capítulo 3 – ABORDAGEM PROPOSTA

A seguir são apresentados os procedimentos metodológicos para os experimentos realizados nesta pesquisa, visando *gerar* abordagem para que o processo de ETL seja integrado aos indicadores de negócios, através de visões materializadas em aplicações de *Data Warehouse*.

3.1 Etapas da Pesquisa

As etapas desta pesquisa foram sistematicamente desenvolvidas no conceito de estudo de caso. Conforme Yin (2005) baseia-se em um conjunto de características associadas ao processo de coleta de dados e estratégias de análise dos mesmos.

Desta forma, as etapas abrangidas nesta pesquisa são:

- Revisão da literatura;
- Coleta de dados (ERP da empresa);
 - Utilização dos indicadores de negócios,
- Experimentação;
 - Preparação dos dados coletados, através da integração do processo de ETL aos indicadores de negócios e de visões materializadas,
- Análise de resultados obtidos na fase de experimentação.

A revisão literária desse trabalho aconteceu em dois períodos, *a priori*, foi realizado um breve levantamento bibliográfico (**capítulo 2**), que teve como objetivo principal apresentar um discurso sobre as temáticas envolvidas, para dar embasamento teórico na discussão dos conceitos e tecnologias utilizados no contexto do trabalho.

O segundo momento da revisão literária foi dado ênfase na apresentação de trabalhos relacionados (**seção 2.5**). A seleção dos trabalhos teve como critério de aceitar apenas pesquisas que envolvem-se contribuição de visão materializada em *Data Warehouse* e utilização de indicadores de negócios nesse processo.

Após, a conclusão desta atividade, as etapas de coleta de dados, seleção dos indicadores de negócios (**anexo 1**), elaboração das visões materializadas (**apêndice**) que serviram de base para implementação da nova abordagem (**seção 3.4**), experimentação (**seção 4.3**) e análise de resultados obtidos (**capítulo 4**).

3.2 Estudo de Caso

A base de dados e os indicadores chaves de desempenho que serviram para implementação da abordagem proposta foram compartilhados pela Escola Crescimento que é uma instituição de ensino de médio porte, foi criada em 1985 com a finalidade de promover educação básica em São Luís, estado do Maranhão. Atualmente é uma das maiores escolas privada da capital. Conta com cerca de 287 (duzentos e oitenta e sete) colaboradores e aproximadamente mais 2.000 (dois mil) alunos, possui certificação de qualidade ISO 9001:2000. Em 2009, a instituição foi recertificada com base na revisão mais atual da mesma norma, a ISO 9001:2008.

O setor que foi objeto de estudo: direção administrativa financeira, esse engloba várias funções, por exemplo, consiste em planejar, organizar, dirigir e controlar as atividades financeiras da empresa, fixar políticas de ação acompanhando seu desenvolvimento, para assegurar o cumprimento dos objetivos e metas estabelecidos, realizar o gerenciamento completo da área administrativa e financeira da empresa, contemplando as atividades de planejamento financeiro, contas a pagar e contas a receber, cobrança, compras administrativas, coordenar as atividades da tesouraria e da controladoria, analisar e acompanhar as execuções orçamentárias, gerir as áreas contábil, financeira e fiscal, realizar análise e apuração de impostos, acompanhar rotinas fiscais, contábil, manter relacionamento com bancos e execução das operações financeiras, dar suporte às áreas de negócios, com a criação e análise de relatórios gerenciais e indicadores de performance, realizar estudos de viabilidade econômica para novos projetos e produtos.

Dentre os vários indicadores e controles desta área, focar-se-á apenas nos que serão objetos de estudo da pesquisa (**anexo 1**). Todos os indicadores tem como base uma estrutura de dados (**anexo 2**), proposta por Neely (2002), vide **Tabela 3**.

Tabela 3- Estrutura de dados de indicadores

Elemento	Descrição
Medida (nome)	O título do indicador. Um bom título deve ser autoexplicativo
Propósito (objetivo)	O motivo pelo qual alguma coisa está sendo medido. Se não há motivo, não há porque ser medido.
Referente a	O objetivo de negócio propriamente dito. Se o indicador não estiver relacionado a um objetivo de negócio, não há porque

	ser medido.
Alvo	Um fim e uma meta específica na qual a organização necessita alcançar dentro de um determinado período de tempo.
Regra	A norma e os padrões pelos quais as pessoas serão estimuladas caminharem em direção do alvo esperado
Frequência	A frequência em que o indicador deve ser atualizado e apresentado. Exemplo: Mensal.
Quem mede?	A pessoa responsável por coletar e reportar o dado.
Fontes de dados	A especificação do local de onde o dado será extraído. Exemplo: Software CRM integrado ao ERP da companhia
Quem age sobre o dado?	A pessoa responsável por tomar ações corretivas ou preventivas em prol do resultado esperado.
O que eles fazem?	A definição de um processo de gestão sobre a ação a ser tomada, no caso, do indicador apresentado ser aceitável ou não.
Observações	Qualquer característica específica, questões pendentes e problemas específicos que tenha a ver com medida apresentada pelo indicador.
Data e número do indicador	O registro da data e da questão que identifica o indicador

Fonte: Neely, 2002.

3.3 Arquitetura do processo de ETL atual

A **Figura 7** mostra a arquitetura atual do processo de ETL, que descreve os dados da fonte sendo extraídos a partir dos dados operacionais (OLTP) por uma ferramenta de ETL, onde os mesmos são transformados e carregados para o *Data Warehouse*. Necessário observar que é possível encontrar diversas ferramentas de ETL já desenvolvidas no mercado, tais como: *open-source* (*Apatar*, *Pentaho Community*, *SpagoBI*, *KETL*, *Palo* etc) ou proprietárias (*Oracle Business Intelligence Standard*, *Microsoft Business Intelligence*, *SAP Business Intelligence* e *IBM Cognos*), ainda assim, tem-se a necessidade de criar rotinas de carga de dados para atender determinadas situações que poderão ocorrer.

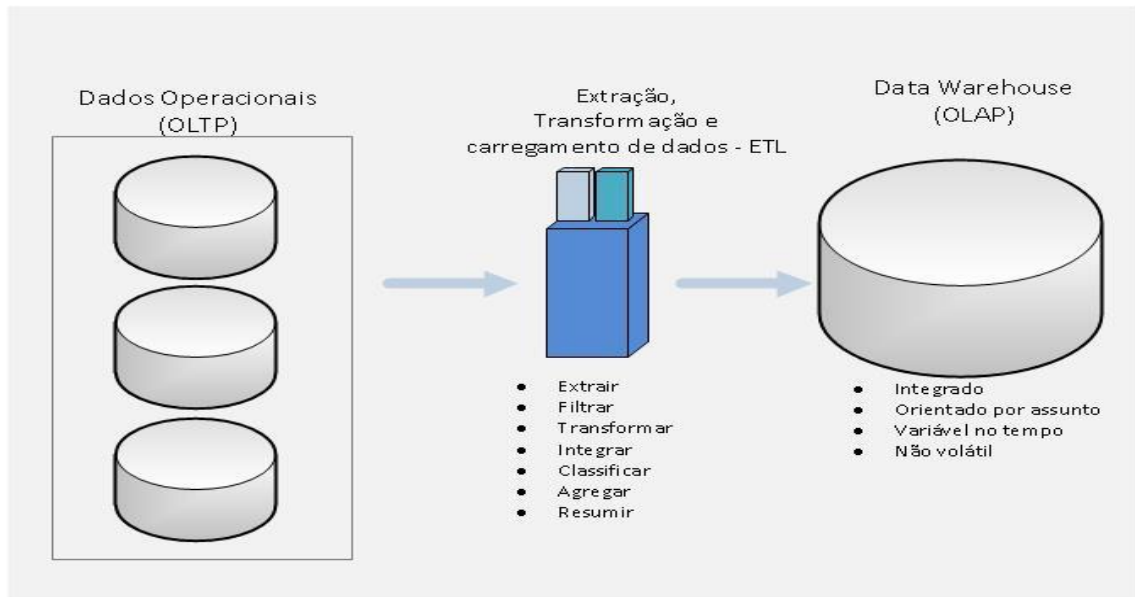


Figura 7- Arquitetura atual do ETL e Aplicação Data Warehouse (Rob e Coronel, 2011)

Nessa arquitetura (**Figura 7**), usa-se no processo inicial, a replicação da base de dados por completo, contribuindo para uma maior alocação de espaço e consequentemente um oneroso tempo de acesso.

3.4 A abordagem proposta

A arquitetura proposta utiliza os indicadores de negócios para extrair dados da base OLTP, para o *Data Warehouse* (OLAP), e esses servem de base para nortear as análises e tomadas de decisões empresariais. Sugere-se a integração destes no processo de ETL (**Figura 8**) de cargas inicial e incremental (através de visões) na base de dados operacional (OLTP), com o objetivo de reduzir o espaço alocado no *Data Warehouse* e, com isso, minimizar o tempo de carga neste processo, contribuindo para reduzir o tempo de processamento de dados para obtenção dos indicadores de negócios.

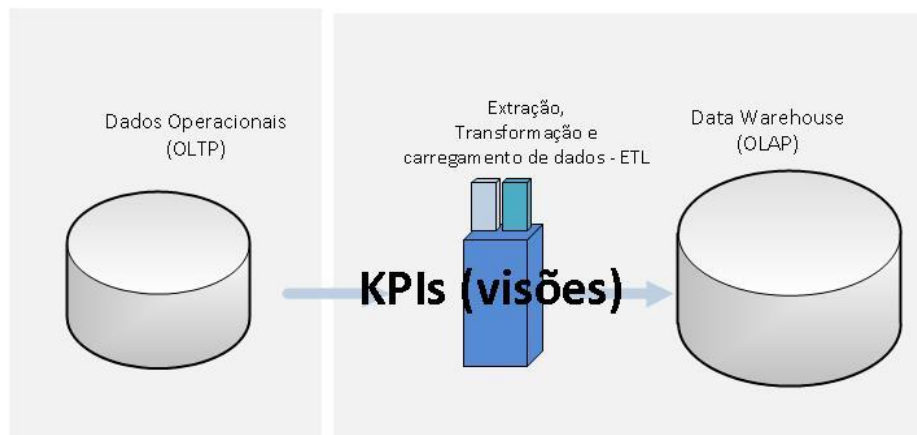


Figura 8- KPI integrado ao processo ETL

3.5 Processo de ETL integrado aos indicadores de negócios

No estudo de caso, o processo de ETL é representado através de scripts SQL. Não se utilizou ferramenta, que auxiliasse na tarefa de extração, transformação, carga e conversão dos dados. Desta forma, obteve-se uma visão mais clara das principais dificuldades relacionadas com estes processos, conforme ilustrada na **Figura 9**.

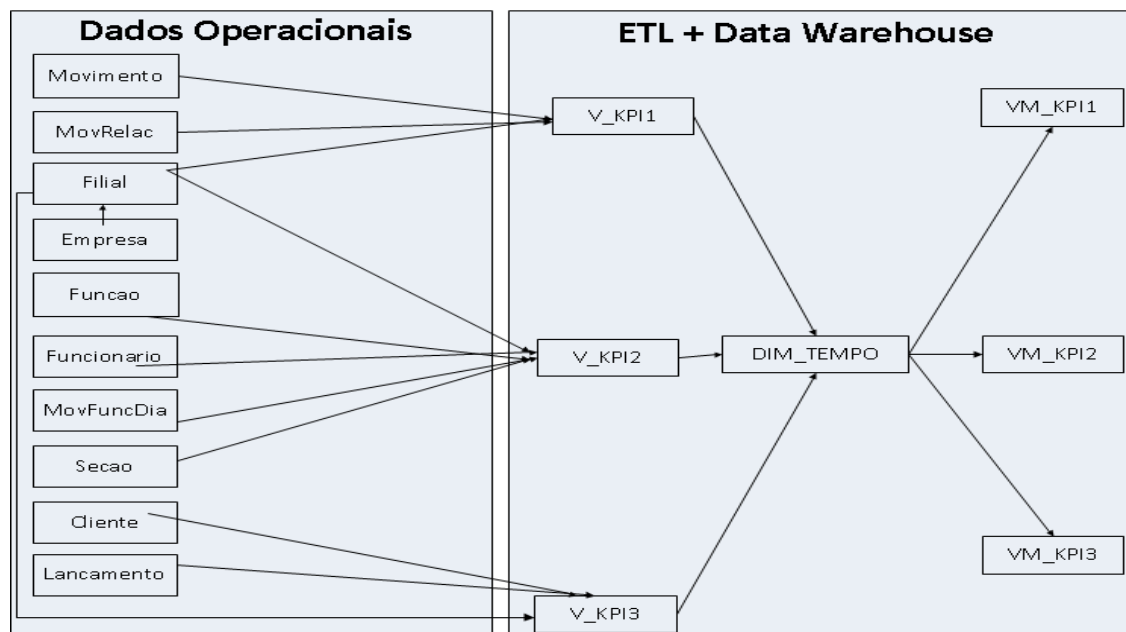


Figura 9- Processo de ETL integrado aos indicadores de negócios

De acordo com a **Figura 9**, os dados são transportados da base de dados operacionais para a base de dados *Data Warehouse* através de visões que são baseadas em indicadores de negócios. Por exemplo, a visão criada denominada V_KPI1 (Eficiência em Entrega de Produtos - Prazo de Atendimento) tem sua base em quatro Tabelas: Empresa, Filial, Movimento e MovRelac e assim sucessivamente para as outras visões. Utilizou-se a mesma nomenclatura no que concerne as tabelas do *Data Warehouse*, pois têm-se como origem as visões da base de dados operacional.

A **Figura 10** mostra a visão baseada no KPI: Eficiência em Entrega de Produtos - Prazo de Atendimento, enquanto que a **Figura 11**, mostra o processo de carga inicial de dados, onde são declaradas as variáveis (linhas: 1 e 2), atribuídos valores (linhas: 3 e 4) e finalmente o uso de consulta (linha: 5) para inclusão dos dados no *Data Warehouse*.

```
CREATE VIEW V_KPI1 (CODEMPRESA, CODFILIAL, EMPRESA, IDMOVENTRADA,
DATAENTRADA, QTDESOLDENTROPRAZO, QTDESOLFORAPRAZO, QTDETOTAL)
AS
SELECT
    E.CODEMPRESA,
    G.CODFILIAL,
    G.NOMEFANTASIA,
    T.IDMOV,
    T.DATAEMISSAO,
    QTDELSOL_DENTROPRAZO=CASE WHEN DATEDIFF(DAY, T.DATAEMISSAO,
S.DATAEMISSAO) <= 3 THEN COUNT(*) ELSE 0 END,
    QTDELSOL_FORAPRAZO=CASE WHEN DATEDIFF(DAY, T.DATAEMISSAO,
S.DATAEMISSAO) > 3 THEN COUNT(*) ELSE 0 END,
    QTDE_TOTAL = COUNT(*)
FROM
    DBO.MOVIMENTO T(NOLOCK)
    INNER JOIN DBO.EMPRESA E(NOLOCK) ON T.CODEMPRESA = E.CODEMPRESA
    INNER JOIN DBO.FILIAL G(NOLOCK) ON E.CODEMPRESA = G.CODEMPRESA
    AND T.CODEMPRESA = G.CODEMPRESA AND T.CODFILIAL = G.CODFILIAL
    INNER JOIN DBO.MOVRELAC R(NOLOCK) ON T.CODEMPRESA =
R.CODEMPORIGEM AND T.IDMOV = R.IDMOVORIGEM
    INNER JOIN DBO.MOVIMENTO S(NOLOCK) ON R.CODEMPORIGEM =
S.CODEMPRESA AND R.IDMOVDESTINO = S.IDMOV
    AND S.CODTMV = '1.2.12'
WHERE
    T.CODTMV = '1.1.70'
GROUP BY
    E.CODEMPRESA,
    G.CODFILIAL,
    G.NOMEFANTASIA,
    T.IDMOV,
    T.DATAEMISSAO,
    S.DATAEMISSAO
```

Figura 10- Visão V_KPI1

```
1. DECLARE @anoInicial Int
2. DECLARE @anoFinal Int
3. SET @anoInicial = 2006
4. SET @anoFinal = 2015
5. SELECT * INTO DWKPI.dbo.V_KPI1 FROM VKPI1 WHERE YEAR(DTEMISSAO_ENTRADA)
BETWEEN @anoInicial AND @anoFinal
```

Figura 11- Processo de Carga Inicial

Após o processo de carga inicial deve-se criar uma tabela *DW_LOAD* (**Figura 12**) na base de dados operacional (somente uma vez), com a relação de todos os KPIs e o último valor de cada identificador principal, para controlar as atualizações.

À medida que estas ocorram a Tabela deverá ser automaticamente atualizada com: o maior valor do identificador principal e a data atual (**Figura 13**). Verifica-se que cada KPI tem um identificador principal, por exemplo, V_KPI1: Movimento.IDMOV.

```
CREATE TABLE dw_Load(
    id_KPI smallint primary key,
    nome_KPI varchar(50),
    id_master_KPI int,
    dataatualiza_KPI datetime)
```

Figura 12- Modelo da Tabela DW_LOAD

```
INSERT INTO dw_Load (id_KPI, nome_KPI, id_master_KPI, dataatualiza_KPI) VALUES
(1, 'V_KPI1- Eficiência em Entrega de Produtos - Prazo de Atendimento', 336607, GETDATE())
```

Figura 13- Inserção do maior valor do identificador principal na Tabela DW_LOAD

3.6 Atualização das Tabelas no *Data Warehouse*

O *script* abaixo deve ser usado nas atualizações das tabelas pertencentes ao *Data Warehouse*. Exemplifica-se apenas na tabela V_KPI1, para adaptá-lo as atualizações das outras tabelas basta apenas alterar o nome da tabela e a condição na tabela DW_LOAD.

A primeira parte do *script*, da linha 1 a linha 11, mostra a declaração de variáveis e a atualização controlada dos registros. É atribuída a variável @idmov o registro da tabela DW_LOAD, conforme a condição. Em seguida a variável @idmasterkpi recebe o através da função **MAX()**, o maior valor lido na visão V_KPI1 que se encontra na base relacional (OLTP).

A segunda parte do *script*, especificamente da linha 12 à linha 27, mostra a carga controlada dos registros. A atualização dos dados é feita somente se o registro na base **DBDW** (OLAP) for menor que o existente na base relacional (OLTP). Na linha 27 ocorre a atualização da tabela **DW_LOAD**. Por último são exibidos todos os registros que foram atualizados na base OLAP.

Script Atualiza_DW

1. - Atualização dos dados nas Tabelas do Data Warehouse (OLAP)
2. -- Exemplo de atualização na Tabela V_KPI1 (OLAP)
3. -- Declaração de variáveis
4. declare @idMov int
5. declare @qtdeRegistros int
6. declare @idmasterkpi int
7. -- Atualiza a variável com o valor do código armazenado na Tabela DW_LOAD que se encontra na base OLTP

```

8. set @idMov = (SELECT id_master_KPI FROM dbo.dw_load where
   id_master_KPI=1)
9. -- Atualiza a variável com o maior valor do código armazenado na
   visão V_KPI1 que está na base OLTP, este valor será utilizado
   para atualizar a Tabela DW_LOAD
10. set @idmasterkpi = (SELECT MAX(idmov) from dbo.V_KPI1)
11. - Copia os registros para a Tabela V_KPI1 que se encontra na
   base OLAP conforme condição estabelecida.
12. INSERT INTO DBDW.dbo.V_KPI1
13. SELECT
14. V.CodEmpresa,
15. V.CodFilial,
16. V.NomeFantasia,
17. V.Idmov,
18. V.DataEmissao,
19. V.DataEmissao,
20. V.QtdeSol_DentroPrazo,
21. V.QtdeSol_ForaPrazo,
22. V.Qtde_Total
23. FROM
24. DBRelac.dbo.V_KPI1 V
25. Where V.Idmov > @idmov
26. Tabela DW_LOAD com o maior registro
27. UPDATE dbo.DW_Load SET id_master_KPI =
   @idmasterkpi,dataatualiza_KPI=GETDATE() WHERE id_master_KPI=1
28. -- Mostra todos os registros que foram copiados para a base
   OLAP na Tabela V_KPI1.
29. SELECT * FROM DBDW.dbo.V_KPI1

```

3.7 Visões materializadas baseadas em indicadores de negócios em aplicações de Data Warehouse

Conforme demonstrado na **Figura 10**, têm-se um exemplo do uso de visões materializadas integradas no processo de ETL e as aplicações OLAP. Neste exemplo, temos 10 tabelas cujo sistema de origem é a base de dados operacional, enquanto o sistema de destino é base de dados *Data Warehouse*, sendo que apenas a tabela `DIM_TEMPO` (**Figura 14**) é adicionada ao esquema warehouse.

```

CREATE TABLE DIM_TEMPO (
  data date,
  ano int,
  mes int,
  dia int,
  mes_nome varchar(9),
  mes_nome_abrev varchar(3),
  bimestre int,
  trimestre int,
  semestre int)

```

Figura 14- Tabela DIM_TEMPO

Nota-se que existem três tabelas V_KPI1, V_KPI2, V_KPI3 (*Data Warehouse*) que foram criadas através de visões estas baseadas em KPIs cujas Tabelas tem sua origem na base de dados operacional. A visão materializada VM_KPI1 é realizada como parte do processo de ETL para transformar dados: EMPRESA, FILIAL, MOVIMENTO e MOVRELAC para V_KPI1. A VM_KPI2 é para transformação de dados: EMPRESA, FILIAL, FUNCAO, FUNCIONARIO, MOVFUNCIDIA e SECAO para V_KPI2. A VM_KPI3 é para transformação de dados: EMPRESA, FILIAL, CLIENTE e LANCAMENTO para V_KPI3. Como exemplo, será demonstrado a definição da visão materializada VM_KPI1 (**Figura 15**).

```
CREATE VIEW VM_KPI1
(CODEMPRESA,CODFILIAL,EMPRESA,ANO,MES,DENTROPRAZO,FORAPRAZO,
QTDELANC)
WITH SCHEMABINDING
AS
SELECT
        V.CODEMPRESA,
        V.CODFILIAL,
        V.EMPRESA,
        T.ANO,
        T.MES,
        SUM(V.LANCPRAZO),
        SUM(V.LANCFORAPRAZO),
        COUNT_BIG(*)

FROM
        dbo.V_KPI1 V
        INNER JOIN dbo.DIM_TEMPO T ON V.DATAVENCIMENTO=T.DATA

GROUP BY
        V.CODEMPRESA, V.CODFILIAL, V.EMPRESA, V.FORNECEDOR,
        T.ANO, T.MES
```

Figura 15- Visão Materializada VM_KPI1

Nesse exemplo, à medida que as alterações se propagarem das tabelas da base de dados operacional para a tabela da base *Data Warehouse*, a atualização ocorrerá sob demanda. As visões materializadas na hierarquia da dependência serão atualizadas automaticamente, portanto estas estarão disponíveis para consulta e análise de dados. Foram desenvolvidas mais visões materializadas com base nos indicadores de negócios (**Apêndice**).

Capítulo 4 – ANÁLISE DOS RESULTADOS

Este capítulo apresenta o ambiente utilizado para testes e análises de resultados de experimentos através de tabelas e gráficos. O objetivo do experimento é avaliar, para um escopo limitado de scripts se o processo de ETL integrado aos KPIs através de visões materializadas, contribuiu para otimização de tempo da carga inicial/incremental, obtenção automática dos indicadores de negócios, e redução de espaço alocado em disco no *Data Warehouse*.

Foi realizada a comparação com o uso da abordagem atual com a proposta através da utilização dos indicadores de negócios nas etapas de carga inicial e incremental para obtenção dos indicadores de negócios.

4.1 Ambiente de execução

- Hardware:
 - Intel Xeon E5620 (7200 MHZ), 24 GB RAM, 1 TB SATA,
- Software:
 - SQL Server 2008 R2,
 - Windows Server 2008 R2 + SP1 64 bits,
- Durante o desenvolvimento do experimento, optou-se pelo SQL Server por ser à base de dados da empresa que compartilhou os dados, porém a abordagem pode ser utilizada na maioria dos SGBDs, que suportem visão materializada;
- Fonte de dados utilizada para objeto de estudo faz parte do sistema corporativo da empresa;
- Os cenários analisados foram dos anos de 2006 a 2015.

4.2 Análises dos resultados

Os testes foram realizados com tomadas de tempos de resposta dos scripts e tamanho da carga inicial. Para cada plano de execução foram realizadas 30(trinta) execuções, desprezando a primeira (o “tempo frio”) e considerando o tempo médio das demais tentativas (O “tempo quente”). As medidas monitoradas durante a execução são o tempo de resposta dos scripts e o tamanho da carga inicial em ambientes distintos (**Tabela 4**).

A comparação dos tempos médios de execução dos scripts nos diferentes cenários foi feita por meio de gráficos que apresentam os tempos de execução de forma normalizada. No primeiro ambiente, a carga executada no banco de dados alvo, utilizando a abordagem atual. No segundo ambiente, a carga é executada no banco de dados alvo utilizando a abordagem proposta.

Tabela 4- Carga inicial e tempo médio de execução

Cenário	Base de Dados (Anos)	Abordagem Atual – Extração Total		Abordagem Proposta - Extração Parcial - KPIs	
		Carga Inicial (MB)	Tempo (s)	Carga Inicial (MB)	Tempo (s)
1	2006-2008	554,3984	13,1394	4,9551	9,5298
2	2006-2010	996,3359	19,8702	100,4500	10,8386
3	2006-2012	1505,9450	27,9576	161,5800	12,2666
4	2006-2014	4714,7270	71,4361	264,5900	14,3225
5	2006-2015	6485,4300	87,9797	292,0200	14,5389

Fonte: Escola Crescimento, 2015

4.3 Experimentos

Esses experimentos podem ser visualizados através dos gráficos abaixo, que a integração do processo de ETL aos KPIS e as visões materializadas são usadas pela carga de trabalho como pode ser observada pela linha da abordagem proposta (AP). Ela representa os benefícios em relação ao tempo de resposta, alocação de espaço em disco e a utilização de visões materializadas que trouxeram para a carga de trabalho.

4.3.1 Experimento 1

A **Figura 16** mostra a evolução do tempo de inserção da carga inicial em relação a massa de dados nos anos de 2006 a 2015. É possível perceber que, abordagem proposta (AP) apresentou desempenhos significativos a abordagem atual (AA), os resultados evidenciam que mesmo em cenários onde houve uma carga de dados mais expressiva, o uso de KPIs para este processo reduziu substancialmente o tempo de resposta em 6,05 vezes melhor em comparação com o resultado da abordagem atual (AA).

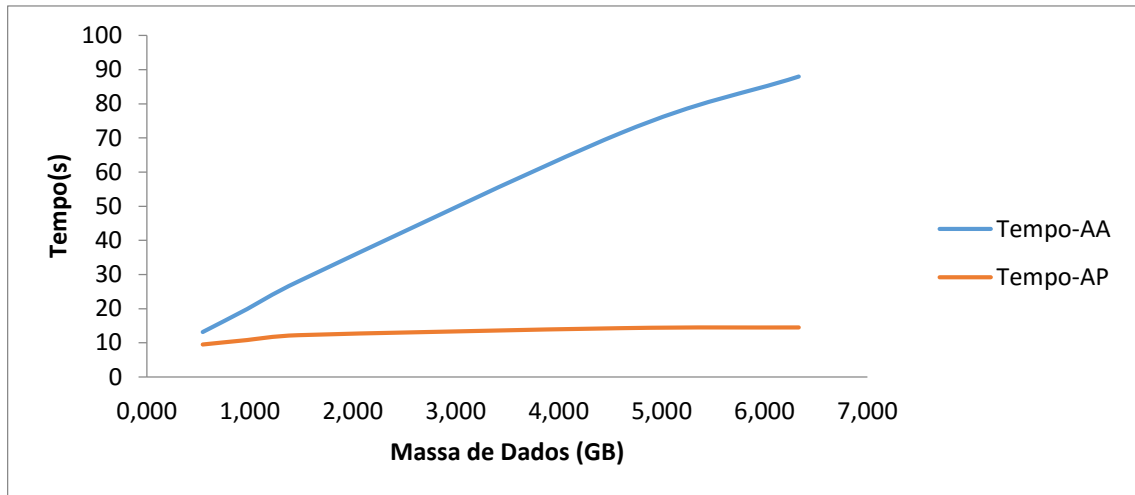


Figura 16 Evolução de tempo para massa de dados da carga inicial

4.3.2 Experimento 2

No experimento 2, a **Figura 17** mostra a evolução da massa de dados nos cenários de 2006 a 2015. Observa-se que apesar do aumento de massa de dados ter se elevado cerca de 22,22 vezes da massa de dados inicial. Assim como, o Experimento 1, o uso de KPI (AP), a redução de espaço alocado em disco foi de aproximadamente cerca de 95,50% em relação a massa de dados inicial.

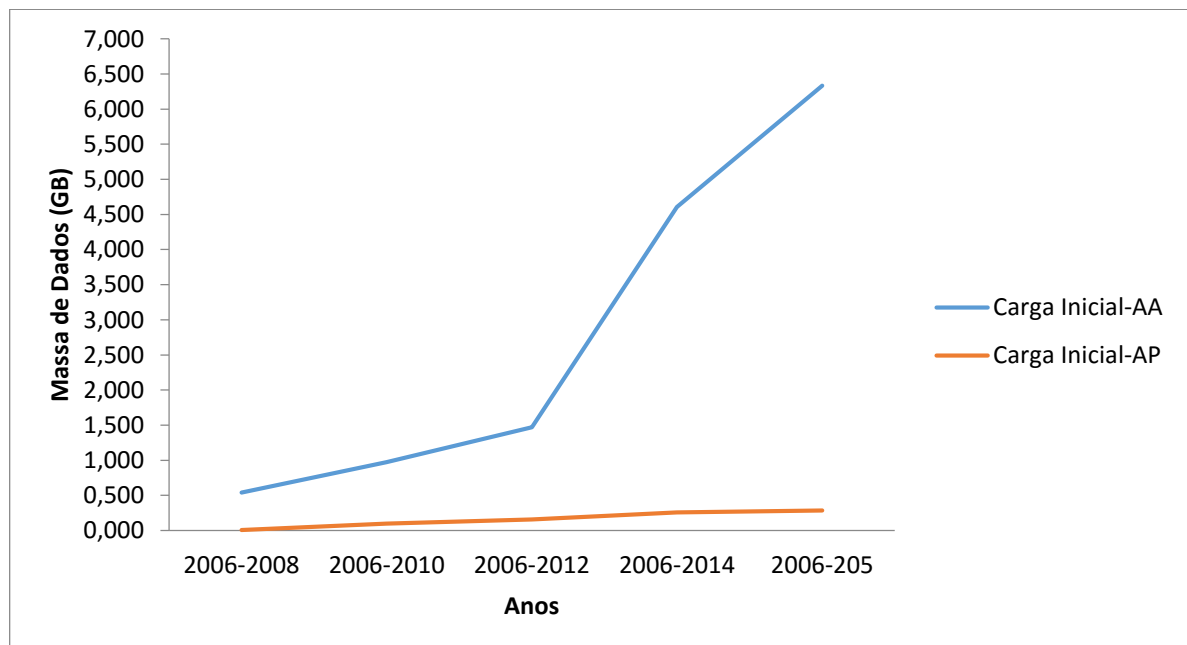


Figura 17- Evolução de espaço alocado em disco no cenário analisado

4.3.3 Experimento 3

Neste experimento, mostra o tempo de resposta para a elaboração, atualização e disponibilização dos indicadores de negócios tanto no processo

manual como automático, através da visão materializada denominada VM_KPI1 (Figura 18), observa-se que, o uso de visões materializadas otimizou o processo em aproximadamente 74,66%. O cenário analisado foi sobre o KPI: Eficiência em Entrega de Produtos - Prazo de Atendimento, com massa de dados total (2006 a 2015).

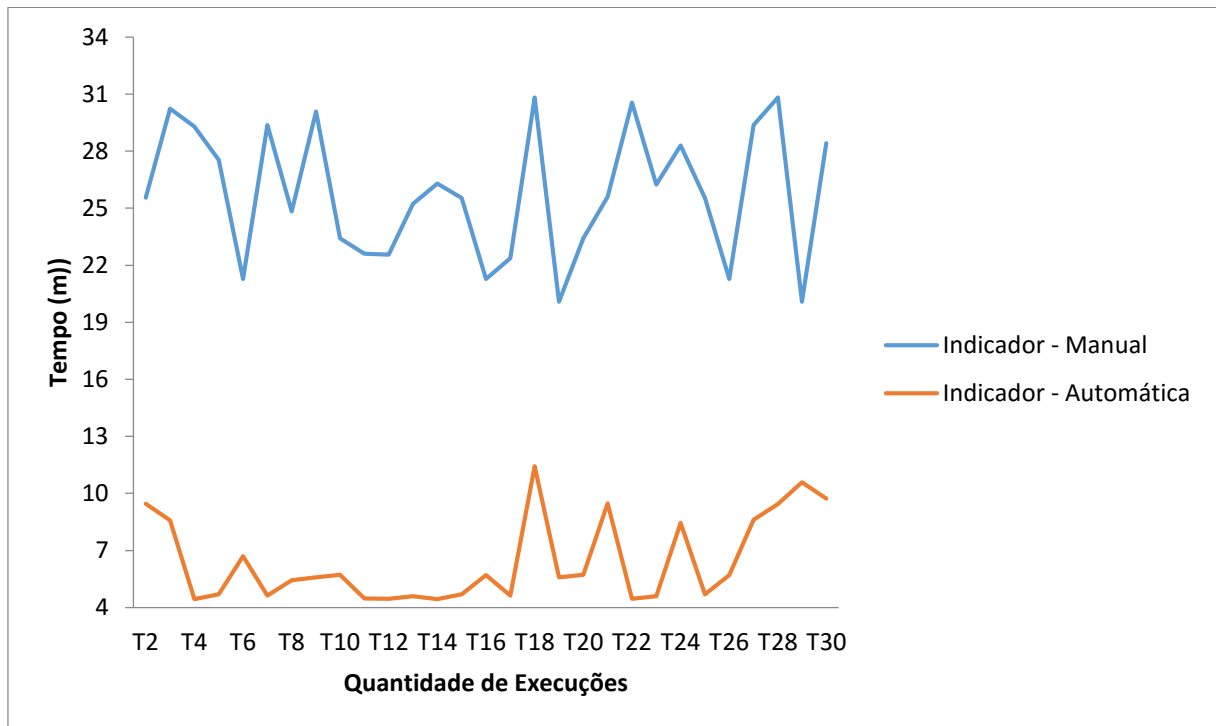


Figura 18- Analogia entre o tempo de resposta V_KPI1 X VM_KPI1

Neste capítulo foram apresentados testes que demonstram a eficácia e a eficiência da abordagem proposta. Os testes realizados com o objetivo de comparar o tempo de acesso, alocação de espaço e tempo de repostas entre as abordagens, mostrou uma melhora significativa no tempo de resposta da carga de trabalho. Com a integração dos indicadores de negócios ao processo de ETL e a construção de visão materializadas o tempo de resposta das consultas melhorou significativamente.

Capítulo 5 – CONCLUSÃO

Esta pesquisa investigou como o processo de ETL, os indicadores de negócios e as visões materializadas, podem auxiliar pequenas e médias empresas em suas tomadas de decisões.

Nesta dissertação foi realizado um levantamento do estado da arte referente a um dos problemas que envolve as pequenas e médias empresas que é a obtenção automática dos indicadores de negócios através de estudo de caso. Para a resolução deste foi proposta uma abordagem baseada no processo de ETL integrado aos indicadores negócios através de visões materializadas em *Data Warehouse* para auxiliar na tomada de decisão. A pesquisa apresentou os passos necessários para utilização destes indicadores neste processo e acredita-se que é eficiente nas alterações do sistema de origem (OLTP), bem como, podem atualizar automaticamente os dados do sistema OLAP, propagando assim para as visões materializadas.

Através dos experimentos realizados buscou-se obter uma abordagem com desempenho bom o suficiente para que fosse capaz de prever que o uso de indicadores de negócios tem sua contribuição importante, no que se refere ao estudo de *Data Warehouse*.

Dentre as contribuições, destacam-se:

- i) Integração do processo de ETL com os KPIs otimizaram a inserção da carga inicial em 6,05 vezes em relação a abordagem atual;
- ii) Redução de espaço alocado em disco, em aproximadamente 95,50% da massa de dados no *Data Warehouse*;
- iii) Obtenção automática dos KPIs através da abordagem proposta e uma otimização no tempo de disponibilização dos mesmos em cerca de 74,66%;
- iv) Confiabilidade das informações;
- v) Independência de funcionários/sistemas na obtenção das informações.

Como proposta para trabalhos futuros, percebe-se que não foram esgotadas todas as fontes de estudo a respeito dessa pesquisa. Pretende-se realizar uma investigação e analogia de implementações de *Data Warehouse* utilizando banco de dados relacionais e *NoSQL (Not only SQL)* visando melhorar aspectos específicos, tais como: alta escalabilidade e disponibilidade, armazenamento e processamento

de consultas eficientes em grandes volumes de dados, alta concorrência na escrita e leitura de dados, e, redução de custos e gerenciamento operacional.

REFERÊNCIAS

BRAVE RIVER SOLUTIONS. **Data Bridge**. Disponível em:

https://www.braveriver.com/_resources/common/userfiles/file/DataBridge-BRS.pdf.

Acesso em 05 mar 2016.

CARVALHO, Andréa Weberling. **Dissertação de Mestrado: Criação automática de visões materializadas em SGBDs relacionais**. Rio de Janeiro: PUC.

Departamento de Informática, 2011.

DATE, C. J. **Introdução a sistemas de banco de dados**. Tradução de Daniel Vieira. 9 ed. Rio de Janeiro: Elsevier, 2003.

DEMING, W. E. **Qualidade e revolução da administração**. Saraiva: Rio de Janeiro, 1990.

DRUCKER, Peter F. **Administrando em tempos de grandes mudanças**. Ed. Pioneira, São Paulo, 2002.

ELMASRI, Ramez; NAVATHE, Sharmkant B. **Sistema de banco de dados**. 4 ed. São Paulo: Pearson Addison Wesley, 2005.

FINGER, Luciane; MUSSI, Ana Paula Toledo Machado. **O uso da visualização de informações para análise de performance: o caso do Sesi Paraná**. Revista Eletrônica Conhecimento Interativo, São José dos Pinhais, PR, v. 8, n. 2, p. 128-144, jul./dez. 2014. 138. Disponível em: <<http://app.fiepr.org.br/revistacientifica/index.php/conhecimentointerativo/article/view/165>>. Acesso em: 13jan 2015.

FUNDAÇÃO NACIONAL DA QUALIDADE (FNQ). **Sistema de indicadores**. Disponível em <<http://www.fnq.org.br/informe-se/publicacoes/e-books/127>>. Acesso em 10 nov 2015.

GOODYEAR, Mark, RYAN, Hugh W., SARGENT, Scott R., et al. **Netcentric and Client/Server Computing – a practical guide**. Andersen Consulting. Boca Raton: Auerbach, 1999.

GUPTA, A.; MUMICK, I. S. **Maintenance of Materialized Views: Problems, Techniques, and Applications**. In IEEE Data Engineering Bulletin, Vol. 18, N.2, Junho 1995.

HARRINGTON, H. James. **Aperfeiçoando processos empresariais**. São Paulo: Makron Books, 1993.

INMON, William H. **Como construir o Data Warehouse**. 2 ed. Rio de Janeiro: Campus, 1997.

INMON, William H.; CHUCK, Kelley. **The Twelve Rules of Data Warehouse for a Client/Server World**. *Data Management Review*. vol. 4(5), May 1994, pp. 6-16.

ITALIANO, Isabel Cristina (2000). **O uso de visões materializadas em data warehouse**. Disponível em: < <http://www.ime.usp.br/~am/5701/isabel-relat.pdf>>. Acesso em 10 jul 2015.

KAPLAN, R. S.; NORTON, D. P. **A estratégia em ação: balanced scorecard**. 4 ed. Rio de Janeiro: Campus, 1997.

KIMBALL, et al. **The Data Warehouse Lifecycle Toolkit**. Wiley, 1998. ISBN 0471255475.

LINGLE, John H.; SCHIEMANN, William A. (1996). **From balanced scorecard to strategy gauge: is measurement worth it?** *Management Review*, March, vol(85).

MOREIRA, E. **Proposta de uma sistemática para o alinhamento das ações operacionais aos objetivos estratégicos, em uma gestão orientada por indicadores de desempenho**. Tese (Doutorado em Engenharia de Produção) - Programa de Pós-Graduação em Engenharia de Produção, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis: 2002.

NEELY, Andy; BOURNE, Mike; MILLS John; PLATTS, Ken; RICHARDS, Huw. **Strategy and Performance: Getting the measure of your business**. United Kingdon: Cambridge University Press, first published 2002.

NEVES, Ricardo Alexandre (2009). **Automatização do processo de criação de visões para modelagem de data warehouse**. Dissertação de Mestrado em Ciência da Computação, da Faculdade de Ciências Exatas e da Natureza, da Universidade

Metodista de Piracicaba – UNIMEP. Disponível em:

<https://www.unimep.br/phpg/bibdig/pdfs/2006/FFCKBCSLVBMR.pdf>

PARMENTER, David. **Key performance indicators: developing, implementing, and using winning KPIS**. John Wiley & Sons, Inc., Hoboken. New Jersey, 2007.

PRIMAK, Fabio Vinicius da Silva. **Decisões com B.I. (Business Intelligence)**. Rio de Janeiro: Editora Ciência Moderna Ltda, 2008.

QUASS, Dallen Wendell (1997). **Materialized views in data warehouses**. Tese de doutorado do Departamento de Ciência da Computação da Stanford University. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=CBD0D66E9B8A289E470CE19FF2EAF2BF?doi=10.1.1.82.234&rep=rep1&type=pdf>. Acesso em Abr 2015.

RAMAKRISHNAN, Raghu; GEHRKE, Johannes. **Sistemas de Gerenciamento de Banco de Dados**. 3 ed. São Paulo: McGraw-Hill, 2008. 912p.

ROB, Peter; CORONEL, Carlos. **Sistema de Banco de Dados: Projeto, Implementação e Administração**. 8 ed. São Paulo: Cengage Learning, 2011.

SILBERSCHATZ, Abraham; Korth, Henry F.; Sudarshan, S. **Sistema de banco de dados**. 5 ed. Tradução de Daniel Vieira. Rio de Janeiro: Elsevier, 2006.

TURBAN, E. et al. **Business Intelligence: um enfoque gerencial para a inteligência do negócio**. Porto Alegre: Bookman, 2009.

YIN, Robert K. **Estudo de caso – Planejamento e métodos**. 3 ed. Bookman. Porto Alegre, 2005.

APÊNDICE

Apêndice - 1

```

1 CREATE VIEW V_KPI_EFETIVIDADE_COMPRA(nomefantasia, ano, mes, qtde)
2 WITH SCHEMABINDING AS
3 SELECT
4 Filial.nomefantasia,
5 Year(entrada.dataemissao),
6 Month(entrada.dataemissao),
7 Count(*)
8 FROM
9 Movimento entrada (nolock)
10 INNER JOINmovrelac (nolock) on entrada.codcoligada = Movrelac.codcolorigem and
11 entrada.idmov = Movrelac.idmovorigem
12 INNER JOINmovimento saida(nolock) on saida.codcoligada = Movrelac.codcoldestino
13 and saida.idmov=movrelac.idmovdestino
14 INNER JOINfilial (nolock) on entrada.codcoligada=Filial.codcoligada and
15 entrada.codfilial=filial.codfilial
16 WHERE
17 Entrada.codcoligada= :CODCOLIGA
18 And entrada.codtmv='1.1.70'
19 And (entrada.dataemissao)>= :DATAEMISSAO
20 And saida.codtmv='1.2.12'
21 GROUP BY
22 Filial.nomefantasia,
23 Entrada.dataemissao
24 HAVING
25 Datediff(day,entrada.dataemissao,saida.dataemissao)<=5

```

Figura 19 Visão Materializa - KPI Eficiência em Entrega de Produtos - Prazo de Atendimento

Apêndice - 2

```

1 CREATE VIEW V_KPI_ABSENTEISMO_MENSALISTA(nomefantasia, segmento, data, atraso,
2 falta, abono, extraexecutado, extraautorizado, v_atraso, v_falta, v_abono, v_extraexecutado,v_extraautorizado)
3 WITH SCHEMABINDING AS
4 SELECT
5 Filial.nomefantasia,Secao.descricao,Movdiafun.data,
6 Sum(atraso),
7 Sum(atraso),
8 Sum(abono),
9 Sum(extraexecutado),
10 Sum(extraautorizado),
11 Cast((sum(movdiafun.atraso)/60) as varchar(5))+':'+
12 Cast((sum(movdiafun.atraso)%60) as varchar(5)),
13 Cast((sum(movdiafun.falta)/60) as varchar(5))+':'+
14 Cast((sum(movdiafun.falta)%60) as varchar(4)) ,
15 Cast((sum(movdiafun.abono)/60) as varchar(5)) + ':' +
16 cast((sum(movdiafun.abono)%60) as varchar(5)),
17 cast((sum(movdiafun.extraexecutado)/60) as varchar(5)) + ':' +
18 Cast((sum(movdiafun.extraexecutado)%60) as varchar(5)),
19 Cast((sum(movdiafun.extraautorizado)/60) as varchar(5)) + ':' +
20 cast((sum(movdiafun.extraautorizado)%60) as varchar(5))
21 FROM
22 Funcionario (nolock)
23 INNER JOIN secas (nolock) on funcionario.codcoligada = Secao.codcoligada and funcionario.codsecao=secao.codigo
24 INNER JOIN filial (nolock) on funcionario.codcoligada = Filial.codcoligada and funcionario.codfilial=filial.codfilial
25 INNER JOIN movdiafun on funcionario.codcoligada= Movdiafun.codcoligada and funcionario.chapa=movdiafun.chapa
26 WHERE
27 Funcionario.codcoligada=@EMPRESA) andMovdiafun.data >= @DATAEMISSAO andFuncionario.codrecebimento = 'M'
28 GROUP BY
29 Filial.nomefantasia, Secao.descricao, Movdiafun.data

```

Figura 20 Visão Materializada – KPI Absenteísmo Mensalista

Apêndice - 3

```

1 CREATE VIEW V_KPI_SALDO_HORAS_MENSALISTA(empresa, inicioper, fimper, segmento, compensar, pagar)
2 WITH SCHEMABINDING AS
3 SELECT
4 Filial.nomefantasia, Saldobancohor.inicioper, Saldobancohor.fimper, Secao.descricao,
5 Case when (saldobancohor.extraant -saldobancohor.atrasoant -saldobancohor.faltaant) +
6 (saldobancohor.extraatu - saldobancohor.atrasoatu-saldobancohor.faltaatu) > 0 then
7 Right('00' + convert(varchar(4), abs(((saldobancohor.extraant - saldobancohor.atrasoant-saldobancohor.faltaant) +
8 (saldobancohor.extraatu - saldobancohor.atrasoatu-
9 Saldobancohor.faltaatu))/60)),2)+':' +right('00'+convert(varchar(2),
10 abs(( (saldobancohor.extraant - saldobancohor.atrasoant -saldobancohor.faltaant)
11 + (saldobancohor.extraatu - saldobancohor.atrasoatu -saldobancohor.faltaatu))%60)),2) end,
12 Case when (saldobancohor.extraant-saldobancohor.atrasoant-saldobancohor.faltaant)+
13 (saldobancohor.extraatu-saldobancohor.atrasoatu-saldobancohor.faltaatu) <= 0 then
14 Right('00'+convert(varchar(4), abs(((saldobancohor.extraant - saldobancohor.atrasoant -saldobancohor.faltaant) +
15 (saldobancohor.extraatu - saldobancohor.atrasoatu -saldobancohor.faltaatu))/60)),2)+':' + right('00'+convert(varchar(2),
16 Abs(( (saldobancohor.extraant - saldobancohor.atrasoant-saldobancohor.faltaant) + (saldobancohor.extraatu -
17 Saldobancohor.atrasoatu - saldobancohor.faltaatu))%60)),2) end
18 FROM
19 Saldobancohor (nolock)
20 INNER JOINfuncionario (nolock) on saldobancohor.codcoligada=
21 Funcionario.codcoligada and saldobancohor.chapa = Funcionario.chapa
22 INNER JOIN secao(nolock) on funcionario.codcoligada =Secao.codcoligada and
23 funcionario.codfilial=secao.codfilial andFuncionario.codsecao=secao.codigo
24 INNER JOIN filial(nolock) on secao.codcoligada=filial.codcoligada
25 and secao.codfilial=filial.codfilial
26 WHERE
27 funcionario.codsituacao<>'U' And funcionario.codrecebimento='M'
28 GROUP BY
29 Filial.nomefantasia, Saldobancohor.codcoligada,Saldobancohor.inicioper,
30 Saldobancohor.fimper, Saldobancohor.chapa, Funcionario.codsecao

```

Figura 21 Visão Materializada – KPI Saldo Hora Mensalista

Apêndice - 4

```

1 CREATE VIEWV_KPI_FOLHA_MENSAL
2 (chapa ,segmento, funcao, evento, anocomp, mescomp, valor, provento,desconto)
3 WITHSCHEMABINDING AS
4 SELECT
5 Funcionario.chapa,Secao.descricao,Funcao.nome,Evento.descricao,
6 Funfinanc.anocomp, Funfinanc.mescomp, Funfinanc.valor,
7 (case evento.provdeschase
8 When 'P' then sum(valor)
9 When 'D' then 0.00 end),
10 (case evento.provdeschase
11 When 'D' then sum(valor)
12 When 'P' then 0.00 end)
13 FROM
14 Funcionario (nolock)
15 INNER JOIN Secao (nolock) on funcionario.codcoligada = Secao.codcoligada and
16 funcionario.codsecao=secao.codigo
17 INNER JOIN funcao(nolock) on funcionario.codcoligada = Funcao.codcoligada and
18 funcionario.codfuncao=funcao.codigo
19 INNER JOIN funfinanc (nolock) on funcionario.codcoligada = Funfinanc.codcoligada
20 and funcionario.chapa=funfinanc.chapa
21 INNER JOIN evento (nolock) on funfinanc.codcoligada = Evento.codcoligada and
22 funfinanc.codevento=evento.codigo
23 WHERE
24 (Evento.provdeschase in('P','D')) and (funcionario.codrecebimento in('M','H')) and
25 (funcionario.codtipo in('N','T')) and (funcionario.codsituacao<>'U') and
26 (funfinanc.nroperiodo=2)
27 GROUP BY
28 Funcionario.codcoligada, Funcao.nome,Funcionario.chapa,
29 Secao.descricao, funfinanc.anocomp,funfinanc.mescomp,
30 evento.provdeschase, evento.descricao

```

Figura 22 Visão Materializada – KPI Folha Mensal

Apêndice - 5

```

1 CREATE VIEW v_kpi_baixa_d2(empresa, ano, mês, qtde totalbaixa, qtdebaixa2)
2 WITH SCHEMABINDINGAS
3 SELECT
4     Filial.nomefantasia as coligada,
5     year(lancfin.dataemissao) as ano,
6     month(lancfin.dataemissao) as mes,
7
8     qtdebaixaforaprazo=case when datediff(day,flan.databaixa,flan.datapag) > 2 then count(*) else 0 end,
9     qtdebaixadentroprazo=case when datediff(day,flan.databaixa,flan.datapag) <= 2 then count(*) else 0 end
10
11 FROM
12     Lancfin (nolock)
13     INNER JOIN gfilial on lancfin.codcoligada = gfilial.codcoligada and
14     lancfin.codfilial = gfilial.codfilial
15 WHERE
16     flan.pagrec in(2,3)
17     and year(flan.dataemissao) >= @ANO
18     and flan.codtdo='0002'
19     and flan.statuslan in(1)
20     and flan.valorbaixado > 0
21 GROUP BY
22     gfilial.nomefantasia,
23     year(flan.dataemissao),
24     month(flan.dataemissao),
25     flan.databaixa,flan.datapag

```

Figura 23 Visão Materializada - KPI Baixa Pagamento

ANEXO

Anexo 1: Indicadores e Controles

ÁREA	PROCESSO	TIPO (I/C)	NOME	PERIODICIDADE	COMO	OPERADOR	GESTOR	FINALIDADE	FONTE	RESULTADO ESPERADO	FORMA DE MEDIR	PRAZO
COMPRAS	Compras - Realização de Compras	I	Eficiência em Compras (Prazo de Atendimento)	Mensal		Assistente de Compras	Gerente Financeiro	Mensurar o tempo do processo de compras desde a solicitação até a entrega de material		Dentro do prazo previsto na política de compras	Relatório fornecido pelo sistema ERP da empresa	dia 5 do mês subsequente
PESSOAL	Departamento de Pessoal - Saúde e Segurança	C	Exames Realizados	Semestral		Coordenador do DP	Diretor Financeiro	Controlar cumprimento da exigência legal da realização de exames periódicos		Não haver ocorrências	Relatório fornecido pelo sistema ERP da empresa	Dias 05/07 e 05/01
	Departamento de Pessoal - Execução da Folha de Pagamento	C	Variação real x orçado (folha base) - Proventos	Mensal		Coordenador do DP	Diretor Financeiro	Mensurar variação entre Real x Planejado na FOPAG		Variação inferior a 5%	Total de proventos FOPAG/Folha base	dia 5 do mês subsequente
	Departamento de Pessoal - Execução da Folha de Pagamento	I	Absenteísmo – Mensalistas	Mensal		Coordenador do DP	Diretor Financeiro	Controlar faltas justificadas e não justificadas		Variação inferior a 3%	Horas de faltas/Horas total trabalhadas	dia 5 do mês subsequente
	Departamento de Pessoal - Execução da Folha de Pagamento	I	Saldo de Horas - Mensalistas	Mensal		Coordenador do DP	Diretor Financeiro	Controlar saldo de horas		Variação inferior a 3%	Total de saldo de hroas/ total de horas	dia 5 do mês subsequente
TESOURARIA	Tesouraria - Gestão de Caixa	C	Baixas realizadas em prazo superior a D+2	Mensal	Através do Relatório Gerencial Mensal	Analista Financeiro	Gerente Financeiro	Controlar as baixas de pagamentos garantindo informações fidedignas na posição de contas a pagar e fluxo de caixa		Zero	Relatório fornecido pelo sistema ERP da empresa	dia 5 do mês subsequente

Fonte: Escola Crescimento, 2015

Anexo 2: Estrutura de dados dos indicadores

Todos os indicadores obedecem a seguinte estrutura de dados:

- 1- Documentação
- 2- Tabulação
- 3- Apresentação

Explicar-se-á cada etapa da estrutura de dados, a título de exemplificação foi utilizado o Indicador: Eficácia no Processo de Compras de Materiais

1- Documentação:

✓ Conceitos:

- O indicador do Sistema de Gestão deve ser uma medida positiva do resultado de um processo ou atividade.
- O indicador deve ter uma meta a ser alcançada, definida segundo padrões de comparação com melhores desempenhos anteriores ou de outras organização, assim como deve ter um padrão mínimo de desempenho, estabelecido segundo os mesmos parâmetros da meta, priorizando-se, quando possível, o desempenho anterior.
- O indicador deve ser revisado pelo menos uma vez ao ano, ou quando houver alterações, com relação aos seguintes itens: a meta e o padrão mínimo de desempenho, responsabilidades, métodos, adequação, entre outros aspectos.
- Os controles do Sistema de Gestão são ferramentas para o monitoramento dos processos e, como tal, não precisam ser necessariamente mensuráveis. Quando forem mensuráveis, devem ter limites para a operação dos processos de forma que viabilizem a observação de variações indesejáveis e a tomada de ações.
- Os indicadores e controles são documentados, apurados e apresentados em arquivos EXCEL, cujo nome deve ter a seguinte lei de formação: AAA - T - Nome do indicador ou controle, onde AAA é a sigla da área responsável e T o tipo, se Indicador ou Controle.
- Os arquivos EXCEL que tratam dos indicadores e controles são criados, formatados e alterados apenas pela ACG (Assessoria de Comunicação e Gestão) e pela TI (Tecnologia da Informação), mediante a solicitação da ACG. Qualquer gestor pode solicitar a criação, alteração ou exclusão de um indicador ou controle, de acordo com sua necessidade.

- Os gestores são responsáveis pelo preenchimento dos campos que definem os indicadores e controles, informações necessárias para o desenvolvimento dos instrumentos de apuração/tabulação e de apresentação (geração de gráficos e relatórios).

✓ **Definições:**

- Objetivo da Gestão: Assegurar o crescimento seguro da organização como forma de garantia de continuidade para nossos alunos.
- Processo/Atividade: Gerenciar Aquisições
- Tipo: I (I- Indicador de Desempenho e C- Controle de Desempenho)
- Nome: Eficácia no processo de compras de materiais
- Finalidade: Mensurar o tempo do processo de compras de materiais desde a solicitação até a entrega ao solicitante.
- Resultado esperado: Dentro do prazo previsto na política de compras de 10 dias
- Forma de medir: O índice é medido com base na mensuração do prazo, tomando como referência a data da solicitação em relação a data de entrega ao solicitante.
- Fórmula:
 - $((1 - (\text{Quantidade de compras no prazo} / \text{Quantidade de compras realizadas})) * 100$
- Meta: 95%
- Periodicidade: Mensal
- Padrão: 90%
- Unidade: %
- Fontes: Relatório do Sistema ERP da empresa.
- Apresentação: O índice é apresentado e analisado através de um gráfico de barra simples acompanhado de uma linha de meta para cada unidade de negócio.
- Prazo de Apuração e Análise: O indicador é apurado até o dia 5 do mês subsequente e analisado até o dia 10.
- Responsabilidades:
 - Área – Financeiro – Compras.

- Operador – Assistente Adm. de Compras
- Gestor: Gerência Financeira
- Histórico de Revisões:
 - Data: 01/02/2015
 - Nº: 1
 - Conteúdo: Revisão completa
 - Justificativa: Reestruturação da cadeia de indicadores e controles.

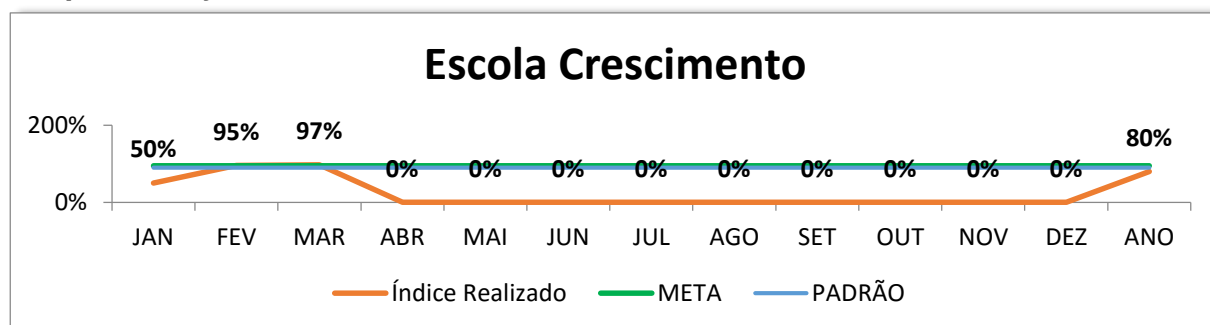
Responsável: Gerente da AGU

2. Tabulação:

		JAN	FEV	MAR	...	ANO
Escola Crescimento	Quantidade no Prazo	432	588	879	...	1899
	Compra Total	858	616	905	...	2379
	Índice Realizado	50%	95%	97%	...	80%

Fonte: Escola Crescimento, 2015

3. Apresentação



Fonte: Escola Crescimento, 2015