

Departamento de Ciências e Tecnologias da Informação

Corporate Performance Management
Aplicação do modelo de Kimball numa empresa de prestação de serviços de
Business Intelligence

Rui Filipe Dionísio Sousa

Dissertação submetida como requisito parcial para obtenção do grau de
Mestre em Engenharia Informática

Ramo de Sistemas de Informação e Gestão do Conhecimento

Orientador:
Professor Henrique O'Neill,
ISCTE-IUL

Setembro, 2011

Folha de rosto



Departamento de Ciências e Tecnologias da Informação

Corporate Performance Management
Aplicação do modelo de Kimball numa empresa de prestação de serviços de
Business Intelligence

Rui Filipe Dionísio Sousa

Dissertação submetida como requisito parcial para obtenção do grau de
Mestre em Engenharia Informática

Ramo de Sistemas de Informação e Gestão do Conhecimento

Orientador:
Professor Henrique O'Neill,
ISCTE-IUL

Setembro, 2011

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Agradecimentos

Gostaria de agradecer ao professor Henrique O'Neill pelo conhecimento que me passou em desenvolver uma dissertação de mestrado e, a todas as experiências académicas que tive o prazer de ter com o professor nos tempos de licenciatura.

Gostaria igualmente de agradecer à BI4ALL a oportunidade de poder realizar um trabalho desta natureza nas suas instalações, agradecendo principalmente ao Dr. Andro Moreira pela possibilidade de arrancar com este projecto, ao Engº Rui Afeiteira e Rui Gorgueira pelo conhecimento técnico que me foi transmitido e utilizado neste trabalho. À Dora Palma e ao Dr. José Oliveira que se mostraram importantíssimos em todas as fases deste trabalho.

Ainda a nível académico, não podia deixar de agradecer aos meus colegas que estão comigo desde a licenciatura, nomeadamente ao Ricardo Barroso e ao Sérgio Carvalho, por todas as experiências pelas quais passámos até à entrega deste trabalho que é a dissertação de mestrado. Muito do que sei hoje é graças a vocês.

A nível familiar tenho que agradecer ao meu pai José, mãe Bárbara e avós Maria e José, por me apoiarem nos tempos em que mais dificuldades tive ao longo deste longo percurso. Todas as condições que tive para realizar este trabalho não seriam possíveis se vocês não tivessem presentes desde o início. Este trabalho é também para vocês e para o avô António e avó Lurdes, que infelizmente, já não estão entre nós mas que mesmo assim merecem sem qualquer dúvida uma grande palavra de agradecimento pelo conhecimento que me passaram ao longo dos anos.

Por último, e não menos importante, gostaria de agradecer à minha namorada Patrícia, que me acompanhou incondicionalmente desde sempre nesta batalha, com o seu apoio e carinho todos os dias. O teu entusiasmo constante comigo, foi sem dúvida, uma das razões de principal sucesso deste trabalho. Por isso, o meu muito obrigado a ti, Patrícia.

Resumo

Esta dissertação surgiu de uma proposta apresentada pela BI4ALL onde se identificou a necessidade em se melhorar o actual sistema de *Corporate Performance Management* já implementado há cerca de 6 anos nessa organização. Dado esse projecto ter sido realizado numa fase muito inicial de vida da BI4ALL, ainda não estavam instanciadas áreas funcionais que hoje existem no dia-a-dia da organização, como a área de controlo e redução de custos. Aliado com a falta de nomenclaturas e boas práticas no desenvolvimento deste tipo de projectos, não foi possível na altura desenvolver uma solução passível de alterações e evoluções ao longo dos anos de vida da organização.

Para atingir este objectivo de reformular de uma forma profunda o sistema já implementado de *Corporate Performance Management* na organização, avançou-se para uma definição de projecto englobando as várias áreas de negócio a serem modeladas no âmbito desta dissertação de mestrado. Irá partir-se para uma solução de *Business Intelligence* desenvolvida completamente de raiz, seguindo todas as metodologias e boas práticas da BI4ALL, modelando e desenvolvendo os modelos dimensionais que vão permitir dar à organização uma nova capacidade de monitorização e análise nas áreas de controlo e redução de custos e, gestão de projectos. Assim, irá resultar um sistema de *Business Intelligence end-to-end*, desde a extração dos dados dos sistemas fonte, até à camada de apresentação onde se encontram os relatórios e o acesso dos dados via *ad-hoc*, contendo todos os *Key Performance Indicators* pedidos pelo negócio. Desta forma, os utilizadores finais da BI4ALL vão conseguir aceder de uma forma muito mais simples e eficiente a um dos activos mais importantes que uma organização necessita de gerir – a informação.

Palavras-chave: *Corporate Performance Management* , *Business Intelligence*, Metodologia, *Key Performance Indicators*

Abstract

This dissertation arose from a proposal by BI4ALL where it was identified the need to improve the current Corporate Performance Management system, already in place for nearly six years in that organization. Since this project has concluded at a very early stage of BI4ALL life, some functional areas were not instantiated and that nowadays are in place at the organization, like the controlling and reducing costs area. Coupled with the lack of classifications and best practices in the development of such projects, it was not possible at that time to develop a capable solution of supporting changes and developments over the years of the organization.

To achieve this goal of a profound overhaul of the Corporate Performance management system already implemented in the organization, it moved to a project encompassing the definition of various business areas to be modeled in this dissertation. It will be broken for a Business Intelligence solution developed completely from scratch, following all the methodologies and best practices of BI4ALL, shapping and developing dimensional models that will allow the organization a new capacity for monitoring and analysis in the areas of controlling and reducing costs, and project management. So, it will lead to an end-to-end business intelligence system, from the extraction of the data from the source systems, to the presentation layer where reports and data access via ad-hoc are located, containing all the Key Performance Indicators requested by the business. Thus, end users of BI4ALL will be able to access in a very simple and efficient way to one of the most important assets that an organization needs to manage – information.

Keywords: Corporate Performance Management , Business Intelligence, methodology, Key Performance Indicators

Índice

| | |
|---|------|
| Folha de rosto | III |
| Agradecimentos..... | V |
| Resumo | VII |
| Abstract..... | VIII |
| 1 Introdução | 21 |
| 1.1 Enquadramento..... | 21 |
| 1.2 Problema | 21 |
| 1.3 Objectivos | 22 |
| 1.4 Motivações para a realização do estudo | 23 |
| 1.5 A estrutura do documento | 24 |
| 2 Estado da arte..... | 25 |
| 2.1 Gestão do desempenho | 25 |
| 2.2 <i>Corporate Performance Management</i> | 27 |
| 2.2.1 Gestão do desempenho dos processos..... | 31 |
| 2.2.2 Análise em tempo real..... | 31 |
| 2.2.3 Integração das aplicações ao longo da organização | 31 |
| 2.3 <i>Business Intelligence</i> | 32 |
| 2.3.1 Os tipos de BI..... | 36 |
| 2.3.2 A base do BI..... | 37 |
| 2.3.3 Architecturas DW..... | 39 |
| 2.3.4 Dimensões..... | 40 |
| 2.3.5 Chaves Substitutas | 40 |
| 2.3.6 Hierarquias..... | 41 |
| 2.3.7 <i>Slowly Changing Dimensions</i> | 41 |
| 2.3.8 Dimensões conformes | 42 |
| 2.3.9 Métricas | 43 |
| 2.3.10 Tipos de tabelas de factos..... | 44 |
| 2.3.11 Modelação dimensional | 44 |
| 2.4 Suporte tecnológico..... | 46 |

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

| | |
|--|----|
| 2.4.1 Microsoft | 47 |
| 2.4.2 IBM | 49 |
| 2.5 Caso de estudo Janseen-Cilag..... | 50 |
| 3 Abordagem metodológica | 53 |
| 3.1 Ciclo de vida de um projecto de sistemas de apoio à decisão | 53 |
| 3.1.1 <i>Program/Project Planning</i> | 54 |
| 3.1.2 <i>Program/Project Management</i> | 55 |
| 3.1.3 <i>Business Requirements Definition</i> | 56 |
| 3.1.4 <i>Technical Architecture Design</i> | 56 |
| 3.1.5 <i>Product Selection and Installation</i> | 56 |
| 3.1.6 <i>Dimensional Modeling</i> | 56 |
| 3.1.7 <i>Physical Design</i> | 57 |
| 3.1.8 <i>ETL Design and Development</i> | 57 |
| 3.1.9 <i>BI Application Design</i> | 57 |
| 3.1.10 <i>BI Application Development</i> | 58 |
| 3.1.11 <i>Maintenance</i> | 58 |
| 3.1.12 <i>Growth</i> | 58 |
| 4 Caso de estudo BI4ALL | 59 |
| 4.1 <i>Program/Project management</i> | 59 |
| 4.1.1 Avaliar a disponibilidade da organização para um projecto de DW/BI | 59 |
| 4.1.2 Endereçar deficiências e determinar os próximos passos..... | 60 |
| 4.2 <i>Business Requirements Definition</i> | 61 |
| 4.2.1 Entrevistas | 62 |
| 4.2.2 Resultados obtidos das entrevistas..... | 62 |
| 4.2.3 <i>Enterprise DW Matrix</i> | 66 |
| 4.3 <i>Technical Architecture Design</i> | 67 |
| 4.4 <i>Dimensional Modeling</i> | 70 |
| 4.4.1 Os quatro passos de Kimball – Área de gestão e controlo de custos .70 | |
| 4.4.2 Diagrama dimensional – Modelo de controlo e redução de custos | 71 |

| | |
|--|-----|
| 4.4.3 Matriz de validação de métricas | 72 |
| 4.4.4 Os quatro passos de Kimball – Modelo de gestão de projectos | 73 |
| 4.4.5 Matriz de validação de métricas | 77 |
| 4.5 <i>Physical Design</i> | 77 |
| 4.5.1 Bases de dados | 77 |
| 4.6 <i>ETL Design & Development</i> | 78 |
| 4.6.1 Configurações nos pacotes | 78 |
| 4.6.2 Extração e transformação | 80 |
| 4.6.3 Carregamento | 84 |
| 4.6.4 Processamento das dimensões e factuais | 87 |
| 4.6.5 Solução Analysis Services – exploração dos dados | 93 |
| 4.7 <i>Designing and Developing Business Intelligence Applications</i> | 97 |
| 4.8 Considerações finais sobre a metodologia adoptada | 102 |
| 5 Conclusões | 103 |
| 6 Recomendações | 105 |
| Bibliografia | 107 |
| Anexos | 111 |

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Índice de tabelas

| | |
|--|-----|
| Tabela 1 - Os três tipos de BI (White 2006)..... | 36 |
| Tabela 2 - Exemplo da aplicação do tipo 2..... | 42 |
| Tabela 3 - Resumo dos indicadores do modelo de controlo e redução de custos | 63 |
| Tabela 4- Resumo dos indicadores do modelo de gestão de projectos | 65 |
| Tabela 5 – <i>Enterprise DW Bus Matrix</i> | 66 |
| Tabela 6 - Matriz de validação de métricas do modelo de controlo e redução de custos | 72 |
| Tabela 7 - Matriz de validação de métricas do modelo de gestão de projectos | 77 |
| Tabela 8 - FCT_CHAMADAS_OPTIMUS – factual do modelo de controlo e redução de custos..... | 112 |
| Tabela 9 - FCT_ACTIVIDADES – factual do modelo de gestão de projectos (sub-modelo de Actividades)..... | 113 |
| Tabela 10 - FCT_ACTIVIDADES_FACTURAVEIS – factual do modelo de gestão de projectos (sub-modelo de Actividades Facturáveis) | 113 |
| Tabela 11 - FCT_TARGETS – factual do modelo de gestão de projectos (sub-modelo de Targets)..... | 114 |
| Tabela 12 - FCT_UTILIZACAO_DE_MAO_DE_OBRA – factual do modelo de gestão de projectos (sub-modelo de Utilização de mão-de-obra) | 114 |
| Tabela 13 - DIM_CLIENTE – dimensão do modelo de gestão de projectos... | 114 |
| Tabela 14 - DIM_COLABORADOR – dimensão do modelo de gestão de projectos e do controlo e redução de custos | 115 |
| Tabela 15 - DIM_LOCAL_TRABALHO – dimensão do modelo de gestão de projectos | 115 |
| Tabela 16 - DIM_PERIODO– dimensão do modelo de controlo e redução de custos | 116 |
| Tabela 17 - DIM_PROJECTO– dimensão do modelo de controlo e redução de custos | 116 |
| Tabela 18 - DIM_TEMPO– dimensão do modelo de gestão de projectos e do controlo e redução de custos | 117 |
| Tabela 19 - DIM_TIPO_CHAMADA– dimensão do modelo de controlo e redução de custos | 117 |

Tabela 20 - DIM_TIPO_COMUNICACAO- dimensão do modelo de controlo e
redução de custos..... 117

Tabela 21 - DIM_TIPO_HORA- dimensão do modelo de gestão de projectos 118

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Índice de figuras

| | |
|--|----|
| Figura 1 - Os três vectores na modelação (Mosimann <i>et al.</i> 2007) | 30 |
| Figura 2 - Tendências do CPM..... | 31 |
| Figura 3 - Exemplo da integração de vários sistemas numa organização..... | 32 |
| Figura 4 - BI <i>framework</i> (Watson, Wixom 2007) | 34 |
| Figura 5 - Independência de várias fontes de dados..... | 34 |
| Figura 6 - Os três níveis do BI (Cardoso 2009) ² , (Quinn 2008)..... | 37 |
| Figura 7 - Fluxo dos dados em BI (Cardoso 2009) ² , (Lukawiecki 2009) | 37 |
| Figura 8 - Back room system architecture model (Kimball <i>et al.</i> 2008)..... | 38 |
| Figura 9 - Abordagem <i>top-down</i> de Inmon (Cardoso 2008), (Caldeira 2008)..... | 39 |
| Figura 10 - Abordagem <i>bottom-up</i> de Ralph Kimball (Cardoso 2008), (Caldeira 2008)..... | 39 |
| Figura 11 - Exemplo de uma hierarquia..... | 41 |
| Figura 12 - Esquema em estrela (BI-INSIDER.COM 2001) | 45 |
| Figura 13 - Exemplo da relação das chaves substitutas (SK) com as chaves estrangeiras (FK) | 45 |
| Figura 14 - Dimensão em “flocos de neve” (Cardoso 2009) ¹ | 46 |
| Figura 15 - Quadrante mágico para plataformas de BI (Sallam <i>et al.</i> 2011) | 47 |
| Figura 16 - Plataforma integrada de BI da Microsoft (Cardoso 2009) ² , (Lukawiecki 2009)] | 48 |
| Figura 17 - Plataforma de BI Microsoft (Cardoso 2009) ² | 48 |
| Figura 18 - Componentes do IBM Cognos BI (IBM 2010)..... | 49 |
| Figura 19 - Arquitectura da solução de BI da Janseen-Cilag | 51 |
| Figura 20 - Diagrama do ciclo de vida de Kimball <i>et al.</i> (2008)..... | 54 |
| Figura 21 - Diagrama das fontes de informação do modelo de controlo e redução de custos..... | 64 |
| Figura 22 - Diagrama das fontes de informação do modelo de gestão de projectos | 66 |
| Figura 23 - Modelo macro da arquitectura desenvolvida | 67 |
| Figura 24 - Sub-sistemas da arquitectura do caso de estudo | 68 |
| Figura 25 - Estrutura de registo da informação ao longo do processo de ETL..... | 69 |
| Figura 26 - Diagrama dimensional do modelo de controlo e redução de custos..... | 71 |

| | |
|--|----|
| Figura 27 - Diagrama dimensional do sub-modelo Actividades - modelo de gestão de projectos | 74 |
| Figura 28 - Diagrama dimensional do sub-modelo Actividades Facturáveis - modelo de gestão de projectos | 75 |
| Figura 29 - Diagrama dimensional do sub-modelo Utilização de mão-de-obra - modelo de gestão de projectos | 76 |
| Figura 30 - Diagrama dimensional do sub-modelo Targets - modelo de gestão de projectos..... | 76 |
| Figura 31 - Configuração de pacotes no SSIS | 79 |
| Figura 32 - Pacote de ET relativo à dimensão Colaborador..... | 80 |
| Figura 33 - Fluxo de dados do pacote de ET da dimensão Colaborador | 81 |
| Figura 34 - Pacote de ET relativo à factual das chamadas do modelo de controlo e redução de custos | 82 |
| Figura 35 - Fluxo de dados (<i>data flow</i>) do pacote de ET da factual das chamadas do modelo de controlo e redução de custos..... | 83 |
| Figura 36 - Pacote de L relativo à dimensão Colaborador | 84 |
| Figura 37 - Fluxo de dados (<i>data flow</i>) do pacote de L da dimensão Colaborador..... | 84 |
| Figura 38 - Pacote de L relativo à factual das chamadas do modelo de controlo e redução de custos | 85 |
| Figura 39 - Fluxo de dados (<i>data flow</i>) do pacote de L da factual das chamadas do modelo de controlo e redução de custos | 86 |
| Figura 40 - Pacote relativo ao processamento global de todo o ETL | 88 |
| Figura 41 - Pacote de processamento das dimensões | 89 |
| Figura 42 - Pacote de processamento das factuais..... | 90 |
| Figura 43 - Pacote de processamento OLAP | 90 |
| Figura 44 - Processamento de dimensões (<i>Analysis Services</i>)..... | 91 |
| Figura 45 - Processamento do cubo do modelo de controlo e redução de custos | 92 |
| Figura 46 - Explorador da solução <i>analysis services</i> do caso de estudo | 93 |
| Figura 47 - <i>Data source view</i> dos modelos | 94 |
| Figura 48 - Estrutura da dimensão tempo..... | 94 |
| Figura 49 - Hierarquia da dimensão tempo..... | 95 |

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

| | |
|--|-----|
| Ilustração 50 - Estrutura do cubo do modelo de gestão de projectos | 96 |
| Figura 51 - Cálculos efectuados no cubo do modelo de gestão de projectos ... | 97 |
| Figura 52 - Exemplo de um relatório geral sobre o modelo de controlo e redução de custos (1/2)..... | 98 |
| Figura 53 - Exemplo de um relatório geral sobre o modelo de controlo e redução de custos (2/2)..... | 99 |
| Figura 54 – Exemplo de um relatório de detalhe sobre um colaborador (1/2) | 100 |
| Figura 55 – Exemplo de um relatório de detalhe sobre um colaborador (2/2) | 100 |
| Figura 56 – Exemplo de um relatório ao nível da chamada de um colaborador | 101 |
| Figura 57 - Navegação sobre os dados do modelo de gestão de projectos usando o IBM Cognos <i>Analysis Studio</i> | 102 |
| Figura 58 – Hierarquia da dimensão Tempo..... | 111 |
| Figura 59 - Hierarquia da dimensão tipo de comunicação..... | 111 |
| Figura 60 - Hierarquia da dimensão tipo de chamada | 112 |
| Figura 61 - Pacotes de ET (extracção e transformação) | 118 |
| Figura 62 - Pacotes de L (carregamento) | 119 |
| Figura 63 - Base de dados da área de teste (<i>staging</i>) | 119 |
| Figura 64 - Base de dados de <i>Config</i> | 120 |
| Figura 65 - Base de dados de DW | 120 |
| Figura 66 - Base de dados de <i>Log</i> | 121 |

Lista de abreviaturas

| | |
|------|---|
| CPM | <i>Corporate Performance Management</i> |
| ETL | Extracção, transformação e carregamento (<i>Extract, Transform and Load</i>) |
| BI | <i>Business Intelligence</i> |
| KPI | <i>Key Performance Indicator</i> |
| DM | <i>Data Mart</i> |
| BAM | <i>Business activity monitoring</i> |
| DW | <i>Data Warehouse</i> |
| ODS | <i>Operational Data Stores</i> |
| OLTP | <i>Online Transaction Processing</i> |
| SSIS | <i>SQL Server Integration Services</i> |
| SSAS | <i>SQL Server Analysis Services</i> |
| SSRS | <i>SQL Server Reporting Services</i> |
| ET | Extracção e transformação |
| CEO | <i>Chief Executive Officer</i> |
| L | <i>Loading</i> |
| ND | Não Disponível |
| DSV | <i>Data Source Views</i> |

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

1 Introdução

1.1 Enquadramento

A área do *corporate performance management* (CPM) tem cada vez mais impacto na posição competitiva das organizações perante o negócio e todo o meio envolvente, desde clientes, concorrência e também ao nível interno.

Tendo em atenção os requisitos necessários para a submissão de dissertações de mestrado no ISCTE-IUL, e o interesse e necessidade apresentada pela BI4ALL em melhorar a sua gestão do desempenho, surgiu a oportunidade de se partir para o desenvolvimento deste trabalho.

Esta proposta surgiu numa forma inicial por parte da BI4ALL ao ISCTE-IUL, com o propósito de aplicar ao nível interno da organização o *know-how* das melhores práticas desenvolvidas pelos colaboradores da BI4ALL nos mais variados projectos em curso e concluídos na sua actual carteira de clientes, sendo também gerado como *output* a dissertação de mestrado.

É pretendido com este trabalho contribuir para melhorar o controlo de gestão das áreas funcionais da BI4ALL, nomeadamente a área de controlo e redução de custos, que contempla toda a componente de comunicações e, a área de controlo de gestão de projectos, que incide na gestão do desempenho dos seus colaboradores ao longo dos vários projectos da organização.

1.2 Problema

Actualmente, as organizações necessitam de gerir a informação que possuem da melhor forma possível, evitando assim que ocorram situações de subaproveitamento deste activo. Muitas vezes, este recurso é frequentemente deixado de parte nas organizações, ignorando o valor que pode trazer ao nível estratégico e não compreendendo como a informação é fulcral para atingirem os seus objectivos.

No passado, foi instanciado nesta organização (BI4ALL) um trabalho de teor semelhante, mas que não contemplou a introdução de novas áreas funcionais ao longo do tempo de vida da mesma, onde nessa altura ainda não existiam as áreas funcionais que hoje existem. Também existe o caso, em que o que se encontra desenvolvido actualmente na organização se encontra obsoleto e, aquando o seu desenvolvimento, não foram contemplados meios para evoluir o que havia sido feito de forma a satisfazer todas as necessidades levantadas pelo negócio no futuro. Sendo assim, aliado com o âmbito definido, o problema pode-se desdobrar da seguinte forma:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

1. A BI4ALL estabeleceu um objectivo de controlo e redução de custos que necessita de ser modelado dimensionalmente. Este objectivo inclui o controlo das chamadas efectuadas pelos colaboradores da organização. Com isto, existe a necessidade de controlar de uma forma mais transparente e imediata todo este processo:
 - 1.1 Confirmar os valores apresentados pela Optimus e propôr condições alternativas ao contracto e,;
 - 1.2 Controlar os custos de comunicações pessoais.
2. A BI4ALL como empresa prestadora de serviços, necessita de gerir a informação associada às horas passadas pelos seus colaboradores nos mais variados clientes e projectos internos. A gestão de projectos que é realizada actualmente, não satisfaz a actual necessidade levantada pelos gestores da BI4ALL em termos de informação que não existe contemplada no modelo dimensional de gestão de projectos em produção;
3. Não só em termos de negócio como em termos mais técnicos, será necessário realizar de raiz todo o desenvolvimento de extracção, transformação e carregamento (ETL) e camada de apresentação, pois o que se encontra implementado não está flexível à mudança e não segue a actual metodologia e boas práticas praticadas pela BI4ALL.

1.3 Objectivos

Como objectivo principal deste trabalho, tem-se a modelação dimensional das áreas funcionais existentes na organização ao longo do tempo e, que foram claramente identificadas na fase de levantamento de requisitos, recorrendo sempre às melhores práticas desde a modelação até à implementação. O uso de novas tecnologias existentes no mercado, aliadas com a aplicação de metodologias de implementação de sistemas de suporte à decisão, permitirá não só um melhor desempenho na fase de implementação do novo sistema, como possibilitará um cenário muito mais transparente na gestão da mudança para o futuro.

Os objectivos desta dissertação podem-se agrupar da seguinte forma:

1. Implementar um novo sistema de *Business Intelligence* (BI) na organização:
 - a) Pretende-se identificar quais os modelos dimensionais que necessitam de revisão e efectuar o levantamento de requisitos junto dos utilizadores finais, para assim contemplar novas perspectivas de análise para o negócio e, a identificação de novos modelos a serem desenvolvidos que permitam à organização gerir o desempenho corporativo de uma forma eficaz e alinhada com as suas necessidades;
2. Desenvolvimento de modelos dimensionais para as áreas de gestão identificadas em 1.a) e de

uma solução baseada em relatórios para acesso aos dados e também para *queries ad-hoc*, dotando assim a empresa de ferramentas que permitam monitorizar e avaliar o seu desempenho corporativo através de *key performance indicators* (KPIs) pedidos pelo negócio;

3. Apurar as vantagens que os modelos implementados irão trazer para a organização, tendo como base todo o desenvolvimento realizado neste caso de estudo real.

Para alcançar estes objectivos, vai-se utilizar as melhores práticas aplicadas pela BI4ALL nos seus clientes, aliado com a metodologia desenvolvida por Kimball *et al.* (2008) para projectos orientados ao desenvolvimento de sistemas de suporte à decisão, onde se parte inicialmente para uma fase de levantamento de requisitos e posterior modelação destas áreas funcionais, o desenho e posterior implementação de uma solução informática, recorrendo-se para isso à tecnologia Microsoft para realizar as operações de ETL e exploração dos dados. Para a camada de apresentação da solução informática vai-se utilizar a tecnologia oferecida pela Cognos, onde irá assentar o desenvolvimento da solução de *reporting* que irá integrar o portal corporativo de forma a permitir a consulta de toda a informação de gestão obtida em resultado deste trabalho, como o acesso via *ad-hoc* à informação, sendo assim as análises mais flexíveis para o utilizador final.

1.4 Motivações para a realização do estudo

A complexidade de várias fontes de dados e a qualidade dos dados que aí estão alojados são os exemplos mais comuns que dificultam todo o processo que está inerente ao desenvolvimento de um *data mart* (DM).

Tratando-se este caso de estudo de um caso real, faz todo o sentido em avançar e implementar uma arquitectura que permita satisfazer todas as necessidades identificadas, eliminando deste modo a lacuna até agora existente na organização em torno desta temática.

Sendo ainda o alvo de estudo uma organização que presta serviços de BI, origina-se desta forma um caso particular de estudo onde se irá resolver um problema real, aumentando a capacidade de tomada de decisão que vai passar a existir para a organização.

A BI4ALL irá sair bastante beneficiada com a realização deste estudo, nomeadamente os gestores de topo que passarão a ter na sua posse informação modelada de uma forma muito mais acessível e rápida e, inclusive, informação que nunca tiveram a possibilidade de ter modelada dimensionalmente.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

O ISCTE-IUL, através do centro ADETTI-IUL, manifesta interesse na forma em que este trabalho irá contribuir com *know-how* necessário para implementar um sistema semelhante ao agora implementado na BI4ALL e em futuros projectos.

A nível pessoal também existem motivações que vão culminar não só em termos académicos como em termos profissionais. Um trabalho nesta área em actual crescimento, num contexto real em que se encontra enquadrado, irá contribuir ao nível da aprendizagem devido a todo o contacto com situações reais e que poderão ser aplicadas mais tarde em futuros projectos pelo autor deste trabalho na BI4ALL.

1.5 A estrutura do documento

Este documento irá seguir a estrutura *standard* adoptada por vários mestrandos, de forma a tornar a sua a sua leitura e interpretações simples.

No primeiro capítulo irá estar reservado à introdução do trabalho contendo o âmbito da dissertação, o problema proposto a ser resolvido, os objectivos delineados, motivações para o trabalho e a descrição da estrutura adoptada para este documento.

No segundo capítulo vai-se expor o estado da arte sobre a temática do CPM, referindo também o que se tem evoluído em termos de BI ao longo dos anos e também outros tópicos relevantes nesta área.

No terceiro capítulo vai-se explanar a abordagem metodológica adoptada para a realização deste trabalho, identificando no que consiste esse método e como se evolui ao longo do mesmo.

No quarto capítulo tem-se a exploração do caso de estudo realizado na BI4ALL, com base na metodologia explorada no terceiro capítulo.

Por último, tem-se os capítulos das conclusões (capítulo cinco) e recomendações (capítulo seis).

2 Estado da arte

2.1 Gestão do desempenho

No panorama em que nos encontramos nos dias de hoje, a conjectura mundial obriga as organizações a se adaptarem e arranjam novas formas de se distinguir da concorrência. Atingir os níveis de desempenho máximos permite fazer uma distinção sobre os seus concorrentes directos, tornando-se um objectivo de qualquer organização que pretenda ser bem sucedida.

Nos primeiros tempos em que o assunto discutido entre as empresas se centrava em gestão do desempenho, não se tinha bem a noção de como as coisas iriam evoluir nas décadas seguintes. A preocupação das organizações em conseguirem avaliar da melhor forma possível os colaboradores da sua organização, fez com que esta temática tivesse a sua ascensão por volta dos anos 80 (Dransfield 2000). Os gestores das grandes empresas tinham como necessidade primária conseguir que os seus colaboradores percebessem o propósito de ali estarem e para o que é que estavam.

A necessidade de gerir os colaboradores como sendo um activo precioso da organização, passou a ser uma preocupação constante dos gestores nos tempos modernos. Cada vez mais o sucesso das organizações que têm como seu negócio principal a prestação de serviços aos seus clientes, está associado não só em simplesmente realizar o serviço requisitado pelo cliente, mas sim como é que esse serviço é feito e por que colaborador é realizado. A realização desse serviço, que pode conter as suas especificidades técnicas e práticas, pode requerer um colaborador com mais experiência e que consiga superar imprevistos não contabilizados pela organização prestadora de serviços no momento em que coloca um colaborador nesse projecto. A gestão dos recursos humanos torna-se um factor crítico que necessita de ser gerido com exactidão. “Com o crescimento do sector dos serviços nos últimos 50 anos, a capacidade de gerir os recursos humanos e as suas habilidades de forma mais eficaz e eficiente continua a ser o factor crítico de sucesso para qualquer empresa de serviços”; “Os prestadores de serviços de tecnologias de informação que normalmente oferecem uma ampla gama de serviços, cada um exigindo recursos com certas capacidades, em mercados caracterizados pela alta e volátil e incerteza das exigências dos clientes” (Hu, Lu 2010).

Ter as pessoas a fazer o que é mais produtivo para a organização e para elas próprias, realizando as suas ambições a nível pessoal e profissional, é algo que precisa de ser igualmente gerido e que se revela complicado de manipular da melhor forma. “Recursos humanos são muito mais complexos para gerir do que máquinas e peças, pedindo novos métodos para melhor capturar e

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

representar essas complexidades.”; “O objectivo final na prestação de serviços é ter as pessoas certas no lugar certo e, na hora certa. (AMR 2006). Um primeiro passo fundamental para alcançar este objectivo é exigir a capacidade de previsão (...)” (Hu, Lu 2010).

Hu, Lu (2010) propõe um método de previsão baseado numa “solução de planeamento integrada”. Preverem que nível de exigência será instanciado num determinado contexto e alocarem de melhor forma os seus activos humanos. “Tais modelos permitem um planeamento mais eficaz em termos de gestão do pessoal em fases mais iniciais do processo de arranque, sendo mais fiável segundo as necessidades dos recursos (...)”, tendo sido desenvolvido um sistema integrado de previsão que analisa dados estatísticos recolhidos de forma a serem reconhecidos padrões que auxiliem a tomada de decisão (Hu, Lu 2010).

Gestão do desempenho não pode ser apenas mais um “tipo” de projecto que é criado num dado instante temporal, em que existe um princípio e um fim, onde se pretende satisfazer uma necessidade que foi colocada pela gestão de topo, julgando-se que é o suficiente e que a organização pode passar a afirmar que pratica gestão do desempenho, passando a ser apenas algo que a partir daí requer apenas manutenção básica. Este conceito necessita de ser cíclico “Gestão do desempenho visa a geração sistemática e controlo do desempenho de uma organização” (Melchert 2004).

Segundo Melchert (2004), gestão do desempenho assenta nas seguintes actividades:

- a) Planeamento do desempenho ;
- b) Medição do desempenho e,;
- c) Recompensar o desempenho.

No primeiro ponto, é o ponto de partida na gestão do desempenho onde se “toma medidas para controlar o desempenho” (Melchert 2004). É essencial para se ter sucesso na implementação deste conceito, ter em consideração o alinhamento entre este ponto do planeamento e a parte operacional, possibilitando que **b)** seja realizado eficientemente e correctamente. Sendo esta ligação da estratégia ao operacional algo muito próximo e dependente entre ambos, devem ser encarados como partes de um todo, e não serem criadas iniciativas distintas que não considerem a priori elementos da estratégia ou da execução operacional, onde o “planeamento estratégico está intimamente ligado à execução operacional” (Melchert 2004). Esta proximidade é justificada com o facto de ser a estratégia que faz com que tudo mude na organização, a nível dos processos e sistemas. Se algo mudar na estratégia e não se reflectir no resto, esse alinhamento não ocorre, comprometendo todo este conceito: “Mudanças de estratégia desencadeiam modificações ao nível dos processos de negócio e sistemas de informação” (Melchert 2004). O último ponto (c))

focaliza-se na parte das recompensas que são atribuídas ao se atingir os objectivos propostos, vindo a reforçar o foco da gestão do desempenho ao nível do colaborador.

Contudo Jingliang *et al.* (2010), identificam um conjunto de problemas que as organizações apresentam na altura de implementarem correctamente o conceito da gestão do desempenho. As organizações necessitam de conciliar as inter-conexões que estão implícitas na organização e saber geri-las como parte de um todo, interiorizando que são variáveis importantíssimas. “(...) Num primeiro momento, gestão do desempenho é separar da gestão da estratégia e, ignorando a orientação e prioridades de desenvolvimento estratégico”. O alinhamento da estratégia com os colaboradores é um requisito crucial para se conseguir obter os níveis de desempenho estipulados, devendo-se ainda ter presente, que gestão do desempenho é uma ferramenta essencial mas que se não for utilizada para o fim ao qual foi destinado, nada de mais-valia é originado. “(...) gestão do desempenho é considerado como avaliação do desempenho, e isso faz com que a comunicação, treino, melhoramentos e uso correcto do desempenho resultem em nada prático nas práticas de gestão do desempenho e, esses inconvenientes do processo de gestão do desempenho são uma das principais razões para um pobre desempenho da empresa” (Jingliang *et al.* 2010).

Para melhor ocorrer o alinhamento dos colaboradores com a organização, conceitos como a missão da empresa foram introduzidos pelos gestores de topo, na busca de alcançar o alinhamento dos colaboradores com a estratégia que se pretendia implementar. Neste contexto, as métricas representavam um papel essencial na avaliação do desempenho da organização mas passariam a ser criticados pela sua orientação de curto-prazo que lhes é atribuída. O conceito de CPM viria então a surgir nos anos seguintes como uma nova alternativa.

É ainda importante ter presente que vários acrónimos podem por vezes confundir a interpretação do investigador. Gerir processos e gerir o desempenho dos mesmos processos são paradigmas e áreas de estudo completamente distintos. “Gestão do desempenho é sinónimo de termos como BPM (*Business Performance Management*), CPM ou EPM (*Enterprise Performance Management*). O importante é não confundir gestão do desempenho com optimização e reengenharia de certos processos de negócio (*Business Process Management*)” (Andonov-Acev *et al.* 2008), (Melchert 2004).

2.2 Corporate Performance Management

Este conceito veio a tomar mais forma por parte das organizações que exigiam uma abordagem diferente da que vinha a ser praticada pelo conceito de gestão do desempenho. Bourne *et al.*

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

(2003), refere que esta abordagem pretende passar de um foco de avaliação do desempenho ao nível do colaborador, para passar a avaliar o desempenho a um nível organizacional, ou se quisermos, corporativo. “Gestão do desempenho é um termo amplamente utilizado dentro dos recursos humanos e tem um significado específico associado com a revisão e gestão do desempenho dos indivíduos. No entanto, há também uma preocupação crescente na avaliação do desempenho e, para isso, medir o desempenho não é o suficiente. Esta medição tem de levar a visão e percepção para a acção” (Bourne *et al.* 2003).

O termo de CPM foi introduzido oficialmente pelos especialistas do grupo Gartner em 2001. “A definição de CPM tem sido consistente desde que os analistas da Gartner *Research* a introduziram em 2001. CPM é um termo genérico que descreve todos os processos, metodologias, métricas e sistemas necessários para medir e gerir o desempenho de uma organização.” (Cognos 2008).

Segundo Melchert (2004), o conceito de CPM assenta em quatro pilares:

- a) Orientação aos processos;
- b) Orientação a metas e métricas;
- c) Metodologia de apoio e;
- d) Suporte das TI.

Em **a)** vem-se reforçar o objectivo do CPM ao ser mais orientado aos processos do que era praticado no conceito de gestão do desempenho. “(...) assegura a integração entre as tarefas de definição da estratégia de uma empresa, implementando-o em processos de negócio” (Melchert 2004). Esta citação vai de encontro com outro autor que também reforça esta integração e automatização das tarefas que compõem os processos de negócio da organização “(...) é otimizar o processo (...) em conjunto com KPIs, de modo a organizá-los de uma forma mais eficaz” (Heß 2006).

Em **b)**, as métricas e objectivos são algo que é definido pela gestão de topo através da estratégia. O alinhamento é assegurado sempre que as métricas definidas para se medir o desempenho corporativo tiverem origem da estratégia que está implementada. “(...) objectivos de negócio têm de ser derivados da estratégia e transformados em métricas para servirem como guias direccionais na medição dos processos” (Melchert 2004).

A metodologia em **c)**, permite fazer a tradução da estratégia para a execução ao nível do próprio processo, ie, da estratégia às métricas, utilizando para isso metodologias como o *balanced scorecard* que permite fazer essa integração. “(...) que fornecem uma estrutura que liga os objectivos estratégicos de métricas para orientar a execução de processos de negócio e, para

medir o seu desempenho.” (Melchert 2004); “(...) que há uma necessidade de aproximação das medidas financeiras e não financeiras que se encaixam dentro de um quadro estratégico” (Bititci 1997).

Por último, as tecnologias de informação fornecem um conjunto de ferramentas que permitem implementar a metodologia em **c)**, permitindo assim realizar uma monitorização das métricas derivadas pela estratégia. “CPM é apoiado por um conjunto de ferramentas de *software* para integrar e analisar o desempenho de dados relevantes, para apoiar a tomada de decisão e para facilitar a comunicação das decisões.” (Melchert 2004).

CPM está aliado às ferramentas de BI que dão o suporte tecnológico mas isto não é suficiente para Charlesworth *et al.* (2004) “(...) opera sobre dados já gerados (histórico)”. Com esta limitação existente nas ferramentas de BI, o CPM propõe que as organizações respondam a três questões fulcrais:

- a)** Como estamos a ir?;
- b)** O que devíamos estar a fazer?;
- c)** Porquê?

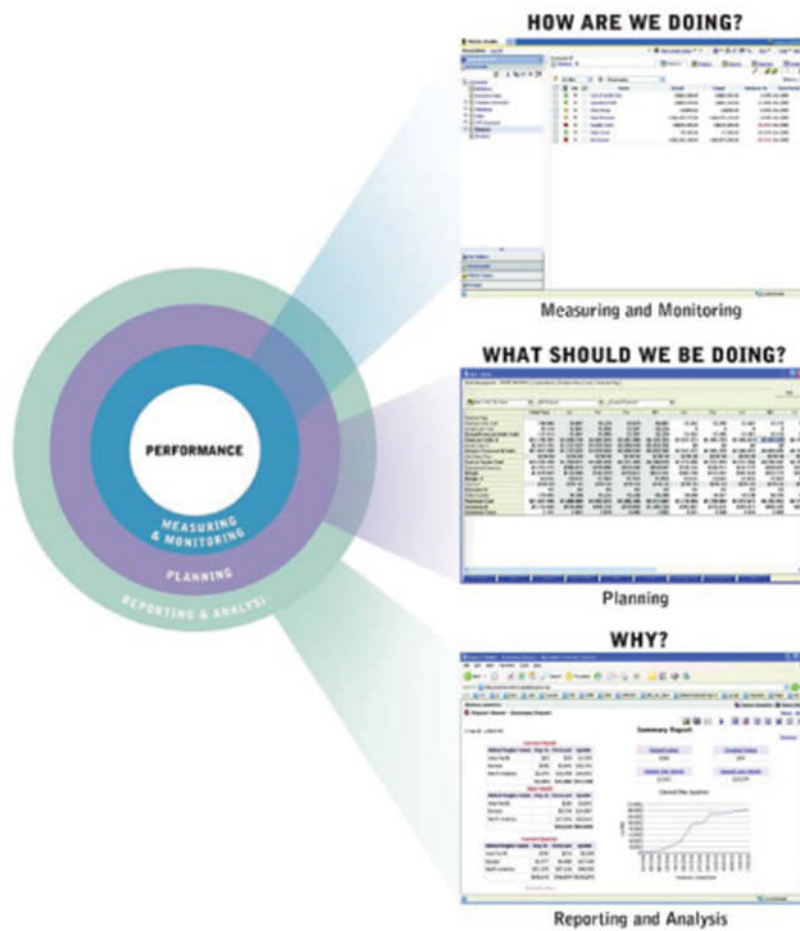


Figura 1 - Os três vectores na modelação (Mosimann *et al.* 2007)

“De acordo com o conceito de CPM, o desenvolvimento de *scorecards*, o BI e as tecnologias de planeamento e consolidação, ajudam a responder a estas três questões que estão interligadas entre si. Saber o que aconteceu, sem saber como, tem pouca importância para a organização. Saber porque é que alguma coisa acontece, sendo incapaz de planear e fazer as mudanças necessárias, é também de pouca utilidade.” (Oliveira *et al.* 2006); “O caminho para uma organização orientada aos processos, tem o CPM como a pedra fundamental onde a empresa se controla permanentemente com base em KPIs, respondendo de uma forma coesa, rápida e flexível às novas exigências e condições de enquadramento, tomando decisões mais cedo.” (Melchert 2004).

Em Melchert (2004) são apresentadas as tendências do CPM e os *enablers* que compõem as mesmas tendências. O conjunto de todos os *enablers* e tendências permitem responder às três perguntas listadas acima e que fazem todo o sentido serem referidas neste documento como forma introdutória.

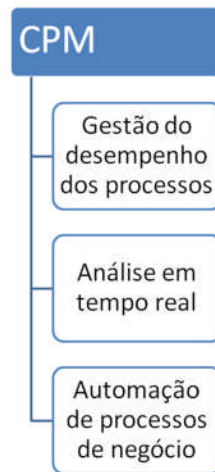


Figura 2 - Tendências do CPM

2.2.1 Gestão do desempenho dos processos

É importante referir que por vezes chega-se a associar a gestão do desempenho dos processos como sendo uma outra designação para CPM. Esta observação nem é de todo errada ou de estranhar, uma vez que CPM é orientado aos processos existentes numa organização: “O caminho para uma organização orientada aos processos, tem a CPM como a pedra fundamental onde a empresa se controla permanentemente com base em indicadores chave de desempenho, respondendo de uma forma coesa, rápida e flexível às novas exigências e condições de enquadramento e, poder tomar decisões mais cedo.” (Heß 2006).

2.2.2 Análise em tempo real

Para este efeito, sistemas de *business activity monitoring* (BAM) foram introduzidos de forma a poderem oferecer informação em tempo real das métricas e indicadores definidos. “O objectivo do BAM é fornecer informações em tempo real sobre o estado e os resultados das diversas operações comerciais, processos e transições” (Kang, Han 2008).

2.2.3 Integração das aplicações ao longo da organização

Este conceito vai de encontro com a necessidade do CPM em integrar os vários sistemas existentes nas organizações. Com esta integração, a informação passará a ser mais facilmente disponibilizada, o que irá possibilitar uma forma de alimentação nos outros sistemas que não ocorria no passado. Um exemplo prático e que vai de encontro ao conceito CPM pode ser o descrito na seguinte imagem, que é baseado numa arquitectura apresentada por Johannesson e Perjons em Melchert (2004):

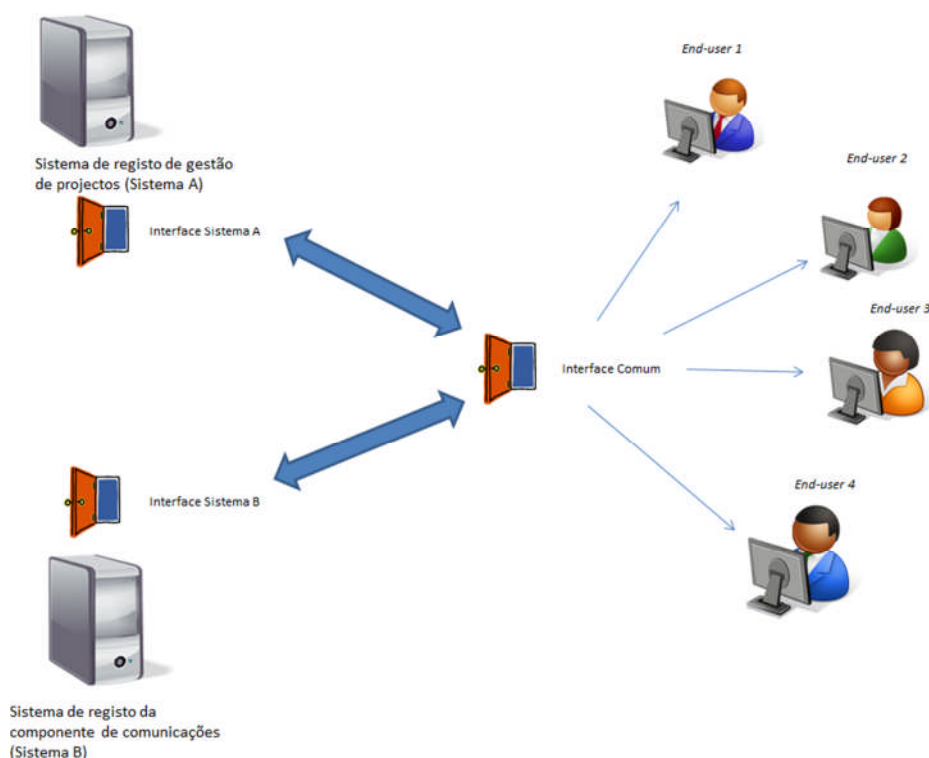


Figura 3 - Exemplo da integração de vários sistemas numa organização

O que se verifica na figura acima, é a forma como uma organização que possui vários sistemas independentes entre si, passam a ter graças ao conceito de CPM um *interface* comum onde os utilizadores acedem da mesma forma tanto à informação armazenada no sistema A como no sistema B, eliminando assim dificuldades que possam existir ao aceder à informação desejada.

A integração dos sistemas numa organização não é assim tão trivial como se pode pensar, podendo existir restrições a nível físico ou a nível de *software*. “As organizações tentam integrar os seus sistemas fonte ao longo de toda a cadeia de valor, enfrentando problemas complexos, tais como: integrar modelos empresariais, interoperabilidade semântica, definição de *interfaces* e padrões de intercâmbio com sistemas antigos, implementando de uma forma segura transacções que ocorrem de uma forma altamente distribuída” (Van den Heuvel *et al.* 2001).

2.3 Business Intelligence

Aparecendo pela primeira vez em meados de 1970, sendo apenas em 1989 instanciado o termo pelo grupo Gartner (Lawton 2006), nem sempre os sistemas de apoio à decisão foram bem aceites pelas organizações como uma mais-valia para o seu negócio. Nos tempos em que surgiu, as análises não eram rápidas o suficiente para que se conseguisse realizar suporte à decisão no tempo desejado, aliada à complexa tecnologia apenas dominada por alguns especialistas na área,

tornando o sistema de BI algo complexo e difícil de se manter (Lawton 2006). Essa realidade viria a mudar ao longo do tempo, tendo-se actualmente “(...) algumas aplicações que possuem interfaces simples o suficiente para uma ampla gama de funcionários a usar. Além disso, algumas implementações funcionam suficientemente rápido para finalidades em tempo real.” (Lawton 2006).

As organizações sentiram a necessidade de automatizar certos processos de negócio, de forma a adquirirem ganhos de desempenho superiores. Essa automatização é então possível através dos sistemas de apoio à decisão, garantindo que o que se mede através desses sistemas vai de acordo com a gestão estratégica implementada na organização. Tais sistemas são os que permitem tornar possível toda a modelação dos processos de negócio, desde a extração, transformação e carregamento dos dados no DW, até às ferramentas que permitem consultar toda essa informação armazenada no DW, facilitando esse processo de extração da informação modelada aos utilizadores finais. “Sistemas de BI combinam a recolha e armazenamento de dados, com a gestão do conhecimento com ferramentas analíticas para apresentar informações internas e de foro competitivo complexas para os gestores com poder de decisão” (Negash 2004), (Watson, Wixom 2007), (White 2006). Os DWs não são nada mais que repositórios de dados, aliados com a capacidade trazida por estes sistemas, fez com que se tornasse o topo da prioridade em termos de projectos nas mais variadas organizações e áreas de actuação, tirando assim o máximo proveito da informação que estas organizações possuem e gerem diariamente (Che *et al.* 2009), tornando essa informação um importante activo estratégico em termos de vantagem competitiva em relação aos seus concorrentes. “(...) o processo de transformar dados em informação e depois em conhecimento” (Golfarelli *et al.* 2004), “Um componente crítico para o sucesso da empresa moderna é a sua capacidade de tirar proveito de toda a informação que tem disponível” (Cody *et al.* 2002). “Um DW é *subject-oriented*, integrado, variante ao longo do tempo e não volátil em termos de dados armazenados para apoio à tomada de decisão (...) a sua arquitectura é dito ser constituída através da integração de dados de múltiplas fontes heterogêneas para dar suporte e/ou consultas *ad-hoc*, relatórios analíticos (...)” (Inmon 1996).

De acordo com o que foi descrito acima, pode-se afirmar que o BI assenta na *framework* apresentada em (Watson, Wixom 2007) e que se encontra na seguinte figura:

1

2

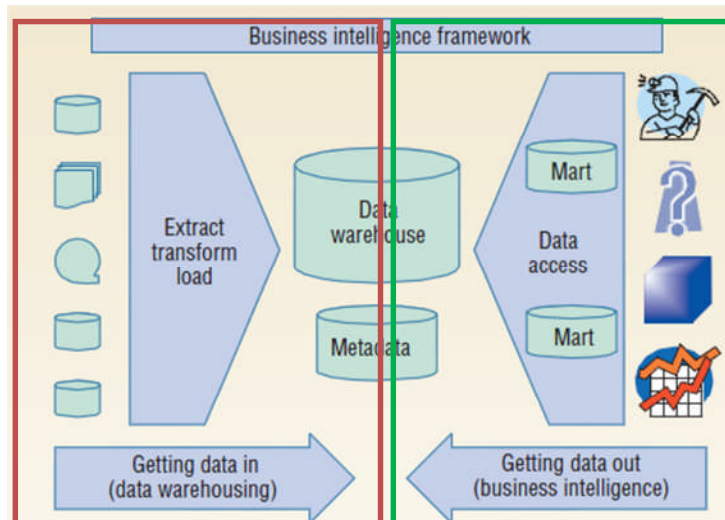


Figura 4 - BI framework (Watson, Wixom 2007)

A ilustração está dividida em dois pontos, sendo o “1” respectivo à parte de ETL, enquanto a “2” está associada à extração dos dados do DW para consulta junto dos utilizadores alvo. É possível constatar que a parte de ETL é associado ao termo *data warehousing*, ou seja, onde os dados são introduzidos no DW (Watson, Wixom 2007), (Cody *et al.* 2002). Esta fase pretende resolver um problema identificado por Mosimann *et al.* (2007): “(...) respostas a partir de fontes desconectadas obscurece o desempenho da organização e dificulta a tomada de decisões.”, ou seja, ao termos várias fontes de dados, torna-se difícil para o utilizador conseguir agrupar toda a informação que necessita, como ilustrado na seguinte figura:

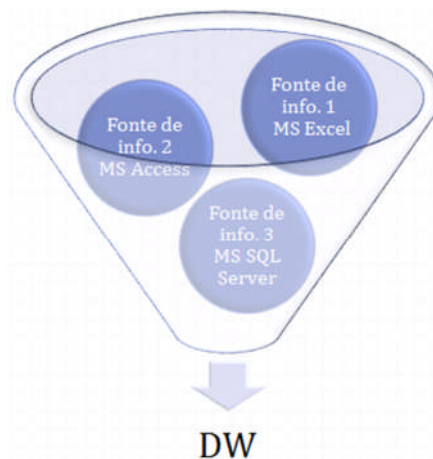


Figura 5 - Independência de várias fontes de dados

Esta figura pretende representar a situação identificada a vermelho na ilustração anterior, em que para além da informação poder estar descentralizada, poderá por vezes assentar em diferentes tecnologias e independentes entre si, dificultando a troca e relacionamento de informação - estes sistemas fonte são também referidos como *Operational Data Stores* (ODS) (Cody *et al.* 2002). Uma outra vantagem ao realizar o *data warehousing* é a da redução dos “(...) custos da infra-

estrutura eliminando processos de extracção de dados redundantes e dados duplicados alojados em DMs independentes ao longo da empresa.” (Watson, Wixom 2007), tornando assim todo este processo de “*getting data in*” conforme ao longo de toda a organização.

Como já referido, na área a verde da figura 4, é onde ocorre o processo de extracção da informação, fazendo uso do DW previamente desenvolvido. Ter os dados no DW não chega para se conseguir tirar partido da informação como um activo estratégico da organização. “Obter os dados dos sistemas fonte permite dar valor limitado à organização; somente quando os utilizadores e as aplicações acederem aos dados e usarem-nos para tomar decisões, é que a organização percebe o real valor do seu DW”. O acto dos utilizadores finais em irem buscar os dados ao DW é referido como *business intelligence*. Chamar *business intelligence* ao acto de extrair os dados do DW tem uma natureza histórica; “O termo *business intelligence* emergiu inicialmente nos anos de 1990 para referir o *reporting* e análise dos dados armazenados no DW. Quando inicialmente apareceu (...) muitas organizações tinham desenvolvido DWs como se fossem arquivos de biblioteca, sem qualquer preocupação em obter os dados e entregá-los ao negócio de uma forma útil.” (Kimball *et al.* 2008).

Esta distinção que é feita nesta *framework*, justifica-se pelo facto de muitos autores da área de BI defenderem o *data warehousing* como o termo que engloba tudo de uma forma geral, e outros que afirmam que *business intelligence* é que é o termo correcto para englobar todos os conceitos (Kimball *et al.* 2008). É então introduzido o termo “*data warehouse/business intelligence (DW/BI)*” – desde a extracção dos sistemas fonte, até à consulta dos dados no DW. A metodologia desenvolvida por Kimball *et al.* 2008, será apresentada num ponto mais à frente (Ciclo de vida de um projecto de sistemas de apoio à decisão).

Ter os dados armazenados no DW não chega para se conseguir transformá-los em informação útil e essencial para a tomada de decisão por parte das organizações. Possuir aplicações que consigam ter a capacidade de realizar *queries* ao DW, isto é, processar os dados que aí estão armazenados, são essenciais para conseguir extrair todo o potencial destes sistemas. Essas aplicações conseguem realizar “suporte à decisão, processamento analítico em tempo real, análise estatística, previsão e exploração de dados”, para as mais variadas finalidades: “traçar perfil de clientes, suporte aos clientes, estudo e segmentação de mercado, a rentabilidade do produto e análise do inventário e a sua distribuição” (Lawton 2006).

2.3.1 Os tipos de BI

É preciso primeiro entender, qual o objectivo da iniciativa de BI a ser implementada na organização antes de se partir para qualquer tipo de desenvolvimento; identificar o âmbito da iniciativa, o tipo de métricas que estão em causa, etc. Para isso, três tipos de BI foram instanciados e podem ser expostos da seguinte forma na tabela abaixo:

| | STRATEGIC BI | TACTICAL BI | OPERATIONAL RIGHT-TIME BI |
|----------------|----------------------------------|--|--|
| Business focus | Achieve long-term business goals | Manage tactical initiatives to achieve strategic goals | Manage and optimize daily business operations |
| Primary users | Executives and business analysts | Senior managers, business analysts, and LOB managers | LOB managers, LOB users, and operational systems |
| Time-frame | Months to years | Days to weeks to months | Intra-day |
| Data | Historical metrics (KPIs) | Historical metrics | Right-time metrics |

Tabela 1 - Os três tipos de BI (White 2006)

Ao nível estratégico pretende-se monitorizar o desempenho, assumindo-se uma perspectiva futura: “alcançar objectivos de negócio a longo termo” (White 2006). Para isso, recorre-se aos KPIs para conseguir traduzir em algo quantificável aquilo que ocorreu no passado, de forma a se identificar o porquê de se estar na actual situação e assim conseguir-se aplicar melhoria contínua nos vários processos, conseguindo-se assim atingir os tais objectivos a longo prazo. (Cardoso 2009)². Este tipo de BI também contempla a comunicação da estratégia de forma a proporcionar a tal melhoria contínua (Cardoso 2009)².

O BI ao nível tático (ou analítico) pretende gerir as iniciativas táticas para atingir os objectivos identificados ao nível estratégico (White 2006). Para isso recorre-se à análise histórica de tendências recorrendo também aos KPIs formulados (Cardoso 2009)².

Por fim, tem-se o BI operacional que se centra “em gerir e otimizar operações do dia-a-dia do negócio” (White 2006). A rapidez com que a informação chega ao utilizador final aliado com o instante definido, permite que as decisões sejam tomadas de uma forma mais rápida, “ajudando as organizações a trabalhar mais ágil e inteligentemente, tornando-se mais competitivas e melhorando a satisfação junto do cliente” (White 2006).

Os três tipos de BI podem ainda ser expostos da seguinte forma, apoiando tudo o que foi referido até ao momento, na seguinte figura:

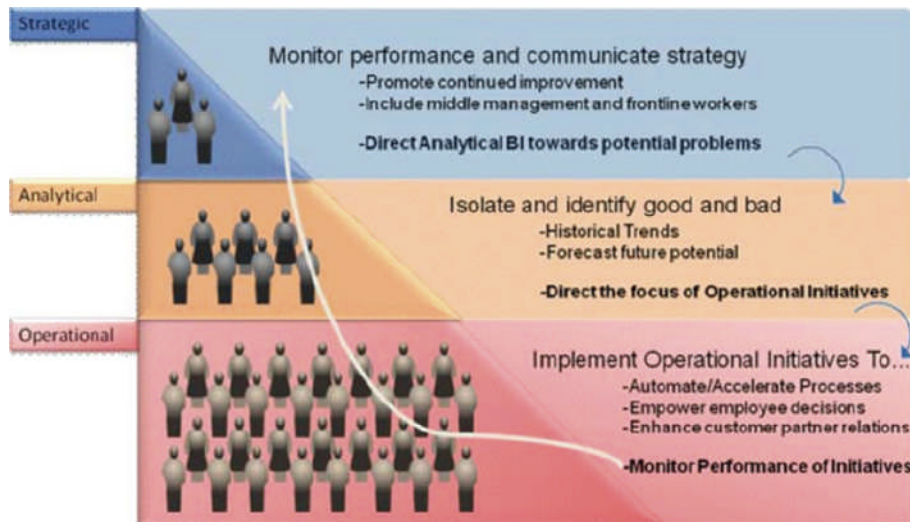


Figura 6 - Os três níveis do BI (Cardoso 2009)², (Quinn 2008)

2.3.2 A base do BI

Antes de se abordar o que compõe o BI em si, é importante primeiro expor a seguinte figura que de uma forma macro, indica os vários sub-tópicos associados ao BI:

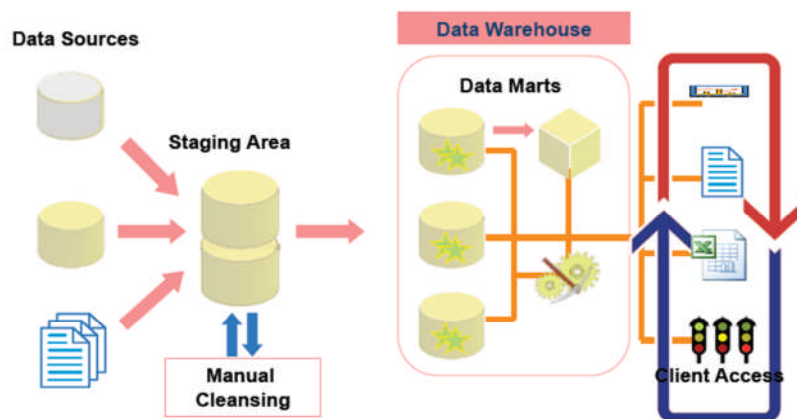


Figura 7 - Fluxo dos dados em BI (Cardoso 2009)², (Lukawiecki 2009)

É possível verificar quatro camadas distintas ao longo de todo o processo:

1. Fontes de dados;
2. *Staging area*;
 - a. Limpeza manual;
3. DW e,;
4. Acesso dos clientes ao DW.

No primeiro ponto (1. Fontes de dados) é onde geralmente existem muitas vezes os chamados silos informacionais. Estes silos são (por vezes) independentes entre si, onde não existe troca de informação entre os vários sistemas existentes, assentando-se muitas vezes em arquiteturas e

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

tecnologias distintas entre si, onde são também conhecidos como sistemas transaccionais. Possuem como características típicas (Cardoso 2009)², (Lukawiecki 2009):

1. Processamento de transacções em tempo real (OLTP);
2. Contêm estruturas de dados optimizadas para modificações;
3. Usualmente fornecem suporte à decisão limitada:
 - a. São habitualmente referidos como sistemas operacionais.

No segundo ponto (2. *Staging area*) e (2. a. Limpeza manual) é onde ocorre a fase de ETL. A fase de extracção está associada à recolha de toda a informação necessária das fontes de dados. Nessa extracção, os dados são “limpos” e armazenados temporariamente nesta *staging area*. Por fim são carregados no DW. Esta fase também poderá ser ilustrada na seguinte figura:

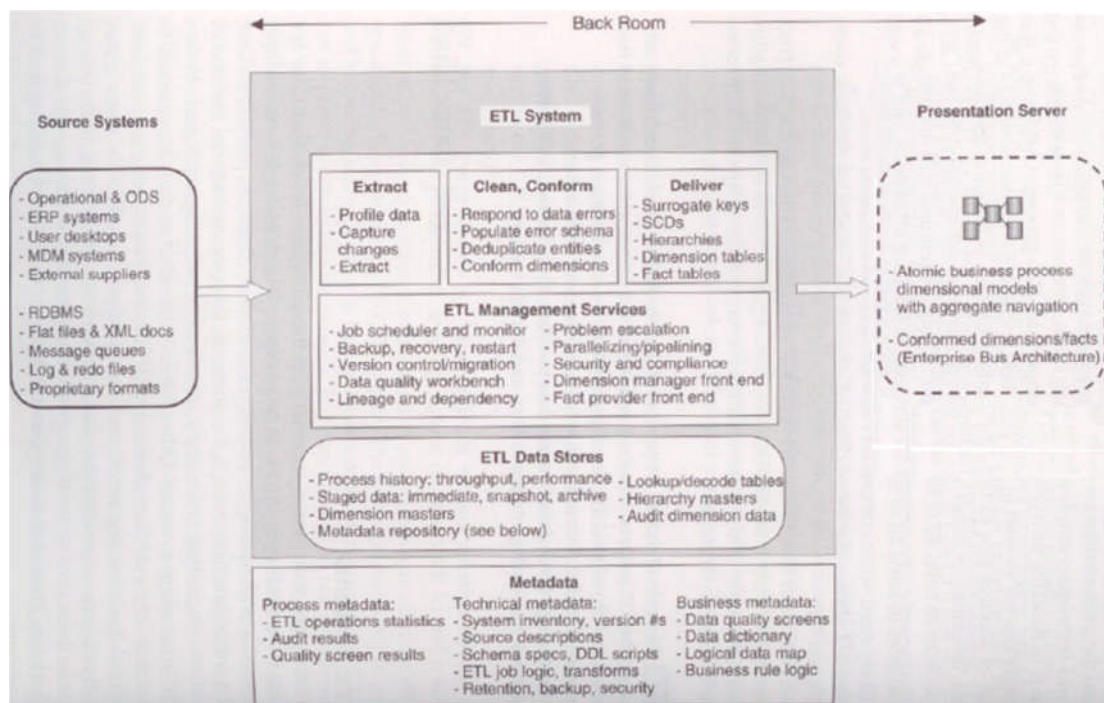


Figura 8 - Back room system architecture model (Kimball et al. 2008)

O terceiro ponto (3. DW) é onde vão estar definidas todas as dimensões e factos. As dimensões são o que permite dar o contexto às análises efectuadas aos factos, ou seja, são entidades que são definidas. Dimensões mais comuns são por exemplo a cliente, produto, tempo, etc. Os factos são as métricas, ou seja, é o que permite quantificar algo que aconteceu no passado. Algumas métricas comuns são o valor da venda, duração de um determinado evento, etc. “Não é muito útil para os decisores do negócio verem apenas uma única métrica agregada representando o total de vendas para todos os produtos ao longo de todos os vendedores para todo o tempo de vida da empresa.” (Cardoso 2009)², (Larson 2009). Importante referir que todas as chaves definidas e utilizadas nos sistemas fonte nunca irão ser as chaves das tabelas presentes no DW – para isso

são criadas as chaves substitutas (*surrogate*) que irão substituir essas chaves vindas da fonte. Mesmo assim, as chaves poderão ser guardadas na mesma no DW, mas nunca como chaves primárias.

O último ponto (4. Acesso dos clientes ao DW) é a camada de apresentação definida para os utilizadores finais acederem a toda a informação contida no DW através de ferramentas de *reporting* e aplicações analíticas.

2.3.3 Arquitecturas DW

Nesta secção, vai-se abordar os pontos de vista *top-down* defendido por W. Inmon e *bottom-up* adoptado por R. Kimball.

A grande diferença entre estas duas abordagens pode ser visualizada nas seguintes figuras 9 e 10:



Figura 9 - Abordagem *top-down* de Inmon (Cardoso 2008), (Caldeira 2008)



Figura 10 - Abordagem *bottom-up* de Ralph Kimball (Cardoso 2008), (Caldeira 2008)

Entre as duas figuras anteriores sobre as abordagens de Inmon e Kimball, consegue-se verificar logo à partida na diferença ao nível do DW. Na abordagem *top-down*, Inmon defende criar-se um DW normalizado, que contenha todos os dados de todos os processos de negócio de uma

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

organização, sendo depois criados DMs que se alimentarão desse DW definido anteriormente. Ou seja, é como se existisse um repositório geral que alimentasse cada DM. Para além destes pontos, esta abordagem tem como características (Cardoso 2008):

- DW como uma aplicação global;
- Desenho complexo e de difícil compreensão para os utilizadores;
- DW normalizado construído com modelo de dados relacional;
- Limitações de desempenho devido ao elevado número de joins necessários às *queries*.

No caso da perspectiva *bottom-up* de Kimball, que foi a adoptada para a realização desta dissertação, o DW é composto por uma série de DMs consolidados logo após a fase de carregamento dos dados realizados no ETL, eliminando-se assim a perspectiva desenvolvida por Inmon. As características desta abordagem são (Cardoso 2008):

- Desenvolvimento progressivo e disponibilização antecipada aos utilizadores;
- DM: estrutura construída com base no modelo de dados dimensional (esquema em estrela);
- Desenho facilmente perceptível aos utilizadores e melhor desempenho das *queries* (pesquisas).

2.3.4 Dimensões

Como já referido, as dimensões são o que dão contexto às métricas contidas nas tabelas de factos, desagregando a informação aí contida. “Reconhecem-se as dimensões e os seus atributos com a palavra "por" (por ano, por produto, por região) que são usados pelas pessoas do negócio quando visualizam um relatório ou uma simples consulta” (Kimball *et al.* 2008), (Cardoso 2009)³. Uma das características mais relevantes numa dimensão centra-se na redundância que as compõe, onde se pode armazenar os códigos que achamos úteis para a caracterizar, mas tendo sempre em atenção que deverá existir um campo mais descritivo. “Não é incomum para uma dimensão ter dezenas de atributos (...) os códigos e abreviações podem ser armazenados como atributos da dimensão, no entanto, é aconselhável incluir um campo descritivo para esses casos” (Kimball *et al.* 2008), (Cardoso 2009)³.

2.3.5 Chaves Substitutas

Sobre as dimensões também é importante referir o conceito de chaves substitutas, cujo propósito é o de deixar que o sistema dependa das chaves primárias que vêm dos sistemas fonte, criando assim as suas próprias chaves substitutas que são simplesmente números inteiros e não contêm

qualquer outro tipo de simbologia utilizada no sistema fonte de onde foram extraídos. O funcionamento prático destas chaves será abordado no ponto **2.3.11 Modelação Dimensional**.

2.3.6 Hierarquias

Em cada dimensão podem-se definir hierarquias, sendo uma das mais utilizadas para exemplos a dimensão tempo. Nós ao termos uma dada métrica, podemos ver o valor atingido, por exemplo, ao nível do mês, ou ao nível do dia. Uma possível disposição de uma hierarquia tempo poderia ser a que está na seguinte figura:



Figura 11 - Exemplo de uma hierarquia

Uma ou mais hierarquias podem coexistir na mesma dimensão, sendo que neste caso o nível mais baixo encontra-se ao nível do dia, e o mais agregado ao nível do ano.

2.3.7 *Slowly Changing Dimensions*

O conceito de *slowly changing dimensions* centra-se na necessidade de se pretender guardar histórico relevante sobre algo ou alguma entidade. Normalmente, uma dimensão não tem os seus dados a alterarem-se com uma ocorrência tão elevada quando comparado com uma tabela de factos. Mas o facto de por vezes nas dimensões essas alterações efectivamente ocorrerem, é necessário possuir uma estratégia de gestão dos dados ao longo do tempo que são alterados. Para isso, Kimball *et al.* (2008) referem três tipos de estratégias que se podem adoptar aquando da alteração de dados numa dimensão:

1. Tipo 1 – Substituir o valor contido no atributo da dimensão;
2. Tipo 2 – Adicionar uma nova linha à dimensão ou,;
3. Tipo 3 – Adicionar um novo atributo à dimensão.

No tipo 1, não existe preocupação em guardar histórico daquele registo. Simplesmente é realizada uma actualização face ao valor antigo, e o novo valor passa a ser o valor de referência. Isto acontece quando para o negócio ter armazenado o valor antigo não traz qualquer valor nas suas análises. Existe uma complicação em termos de implementação deste tipo 1: “Dados pré-agregados com o valor antigo, quando ocorre a actualização ao valor, qualquer sumarização *ad-*

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

hoc realizada e baseada no novo valor não irá corresponder com os dados pré-agregados. Os dados agregados necessitam de ser também actualizados” (Kimball *et al.* 2008).

O tipo 2 é o mais actualizado e, face ao tipo 1, permite armazenar histórico. O modo de funcionamento centra-se em adicionar uma nova linha para cada registo que se pretenda actualizar. Essa nova linha irá manter “o seu identificador primário como um atributo, juntamente com uma coluna contendo a data efectiva do registo, uma coluna com a data de expiração do registo e, uma coluna a identificar qual é a linha actualmente activa” (Kimball *et al.* 2008). Um exemplo deste tipo 2 está contido na seguinte tabela:

| PRODUCT KEY | Product Description | Product Code | Department | Effective Date | Expiration Date | Current Row Ind |
|-------------|---------------------|--------------|-------------------|----------------|-----------------|-----------------|
| 12345 | IntelliKidz 1.0 | ABC99-Z | Education | 2/15/2007 | 5/31/2007 | Not Current |
| 25984 | IntelliKidz 1.0 | ABC99-Z | Stategy | 6/1/2007 | 12/31/2007 | Not Current |
| 34317 | IntelliKidz 1.0 | ABC99-Z | Critical Thinking | 1/1/2008 | 12/31/9999 | Current |

Tabela 2 - Exemplo da aplicação do tipo 2

Por último, o tipo 3 simplesmente insere uma nova linha na dimensão para o novo registo. A vantagem deste face ao do tipo 2 é que permite ter sempre as duas linhas activas, podendo-se realizar análises tanto pelo novo registo, como pelo antigo.

2.3.8 Dimensões conformes

Este conceito foca-se no facto das organizações actualmente serem (na sua maioria) divididas por departamentos. Devido a esta constatação, ter vários departamentos ao longo de uma organização dificulta o processo de levantamento de necessidades no início de vida de um projecto de BI – diferentes departamentos implica, por vezes, necessidades diferentes. Pode fazer sentido para o departamento de vendas ter a dimensão produto, com os atributos que mais sentido fazem a esse departamento, enquanto que para o departamento comercial a modelação realizada nessa dimensão produto para o departamento de vendas não lhes fazer sentido nas suas análises, tendo que existir duas dimensões produto no DW.

As dimensões conformes pretendem eliminar esta redundância de dimensões no DW, existindo apenas uma dimensão (a título de exemplo) produto, que é partilhada por todos os esquemas em estrela e, utilizada por todos os departamentos. Isto permite um alinhamento transversal por toda a organização, contribuindo para um DW muito mais limpo em termos de desenho e na facilidade de manutenção e evolução de todo o sistema de BI.

2.3.9 Métricas

Os factos são também denominados por métricas, onde é possível quantificar um determinado evento que ocorreu. Os exemplos mais típicos encontrados em tabelas factuais são métricas como o *valor total a receber*, *quantidades vendidas*, etc. O conceito de métricas em BI, pode-se desdobrar nos seguintes três tipos (Kimball *et al.* 2008), (Cardoso 2009)³:

- Métricas totalmente aditivas;
 - “Quando faz sentido sumarizar a métrica por todas as dimensões presentes no modelo”.
- Métricas semi-aditivas e,;
 - “Podem ser filtradas por algumas dimensões, mas não todas. Podem ser armazenados nas tabelas de factos, mas com cuidados especiais. Ex: (...) Nível de stock”.
- Métricas não-aditivas.
 - “Não podem ser sumarizadas por qualquer dimensão presente no modelo”.

Depois de se classificar o tipo das métricas em causa, é necessário verificar a granularidade da tabela de factos. A granularidade indica o nível mais baixo de detalhe que cada linha é armazenada na factual (Kimball *et al.* 2008), (Cardoso 2009)³. De acordo com o grão que é definido para essa factual, isso irá determinar quais as dimensões que irão fazer parte do modelo em causa.

“As tabelas de factos devem conter o nível mais baixo, ou seja, o grão mais detalhado possível que é extraído de um processo de negócio” (Kimball *et al.* 2008), (Cardoso 2009)³.

Para se definir o grão deve-se tentar sempre tentar descrever o que é uma linha da tabela de factos, ou seja, o que é que ela representa. Ao se juntar as dimensões é criado o contexto a essa linha da factual. Um exemplo poderá ser: Uma linha da tabela de factos representa a saída de armazém de um produto de uma certa loja, num dado dia, de um certo local, para uma dada concessão. Daqui retira-se que estamos perante uma métrica que será a quantidade saída de armazém, e das dimensões produto, loja, tempo, local e concessão. Isto é algo que é efectuado seguindo a metodologia de Kimball *et al.* (2008).

2.3.10 Tipos de tabelas de factos

Para além do tipo de métricas, também existem tipos de tabelas de factos distintos para se conseguirem reproduzir da melhor forma o negócio que se tenta modelar dimensionalmente.

Para isso, existem três tipos de tabelas de facto (Kimball *et al.* 2008):

1. Transacionais;
2. *Snapshots* periódicas e,;
3. *Snapshots* acumulativas.

A primeira é a mais comum de se encontrar nos modelos. “O grão da tabela é uma linha por transação, ou uma por linha de uma transação (...) num dado ponto no tempo e espaço”. Este tipo de tabelas é reconhecida num modelo dimensional quando tem a si associada várias dimensões. “Sempre que há uma transação, o contexto à sua volta é altamente capturado e registado (...) permitindo monitorizar ao nível mais granular o negócio em causa” (Kimball *et al.* 2008).

O segundo tipo de tabelas de facto são as *snapshots* periódicas, que também são usadas com muita regularidade. Com este tipo de tabelas consegue-se obter fotografias (*snapshots*) do negócio num espaço temporal definido. Esse espaço temporal pode ser definido com uma periodicidade “diária, semanal, ou mensal” (Kimball *et al.* 2008).

O último tipo de tabelas de facto são as *snapshots* acumulativas e, são as que menos se encontram nos modelos dimensionais. A diferença face à do tipo 2, centra-se em que “o intervalo temporal não está definido a priori, mas o processo em causa tem bem definido o início e fim” (Kimball *et al.* 2008). O que isto significa, é que na factual vão-se ter vários pontos temporais, como por exemplo: A encomenda do produto, o envio do produto, chegada do produto, etc. Todas as fases intermédias são sabidas a priori, mas não se sabe se vão acontecer em uma semana, um mês, ou com outra periodicidade qualquer. Num modelo dimensional são facilmente identificáveis, pois a única dimensão presente é a tempo.

2.3.11 Modelação dimensional

Este ponto tem como objectivo expôr as formas possíveis de modelação dimensional postas em equação na realização desta dissertação. Tem-se então as duas formas de modelação; esquema em estrela (*star schema*) e esquema em floco-de-neve (*snow-flake schema*).

O objectivo desta modelação pretende trazer um maior “entendimento para o utilizador” e “maior desempenho na consulta dos dados”, onde são capturadas as “métricas mais importantes para o negócio e respectivas dimensões, pelas quais as métricas são divididas.” (Kimball *et al.*

2008), (Cardoso 2008). Também é referido como útil em casos de certos tipos de *software* que aceitam melhor este tipo de modelação: “Apesar de não ser norma, o esquema em estrela pode ser útil com as ferramentas de *software* correctas. Ao contrário do *design* entidade-relacionamento, o *design* dimensional é totalmente receptiva ao armazenamento redundante de informação.” (Adamson 2010).

Um esquema em estrela de alto nível encontra-se na seguinte figura:

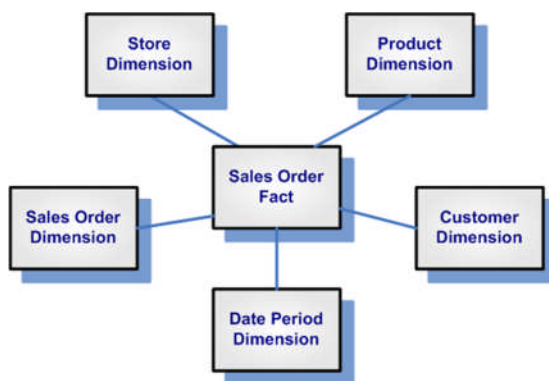


Figura 12 - Esquema em estrela (BI-INSIDER.COM 2001)

Como se pode constatar, no centro do modelo/esquema encontra-se a factual onde estão incluídas as métricas referidas anteriormente. Em redor desta factual encontram-se as várias dimensões pelas quais fazem sentido filtrar a informação contida na factual. As ligações entre a factual e as dimensões são efectuadas através das chaves substitutas criadas na fase de ETL. Na fase de ETL, mais propriamente na criação das dimensões, cada dimensão irá conter uma chave substituta; a tabela de factos irá receber essas chaves substitutas vindas das dimensões como chaves estrangeiras (*foreign keys*). Um diagrama simples contendo essas chaves poderá ser o seguinte:

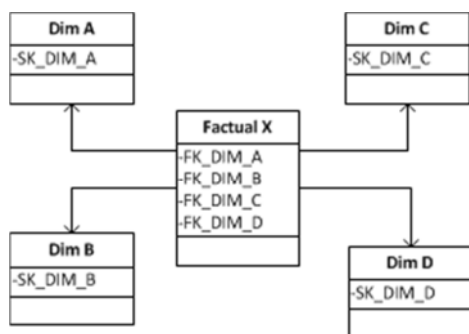


Figura 13 - Exemplo da relação das chaves substitutas (SK) com as chaves estrangeiras (FK)

No esquema em floco-de-neve “cada dimensão normalizada corresponde a uma hierarquia de tabelas.” (Bonifati 2001). Esta abordagem implica que o modelo se vai tornar mais difícil de se compreender quando comparado com o esquema em estrela: “Esta abordagem torna possível

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

reduzir a replicação de dados, mas o diagrama é mais complicado que o tradicional modelo em estrela.” (Bonifati 2001), tal como ilustrado na seguinte figura:

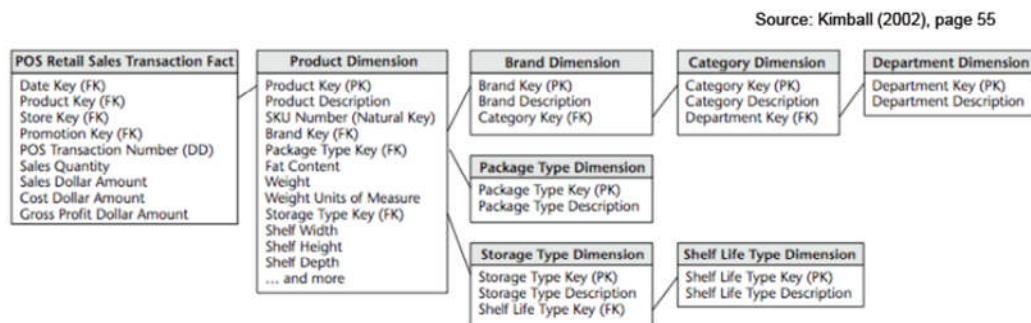


Figura 14 - Dimensão em “flocos de neve” (Cardoso 2009)¹

2.4 Suporte tecnológico

Em BI, actualmente, várias organizações possuem ao seu dispor um leque variado de fornecedores nestas tecnologias, inclusivamente soluções *open-source* que trazem às organizações uma forma de redução de custos dada a actual conjectura mundial que se centra cada vez mais na contenção e, num maior aproveitamento dos activos existentes. “Dado o aumento de interesse em redução de custos expressado nos inquéritos aos clientes, e com o mercado *open-source* em crescimento, era expectável que os mais conhecidos revendedores de *open-source* (Jaspersoft e Pentaho) fossem de encontro com os critérios de aceitação para o quadrante mágico para este ano” (Sallam *et al.* 2011). Portanto, quem tem o objectivo de aplicar e instanciar o BI nas suas organizações, deve ter em conta a oferta existente no mercado, tendo como base os quatro quadrantes mágicos apresentados pela Gartner (Sallam *et al.* 2011). Os quadrantes mágicos apresentam-se na seguinte figura:

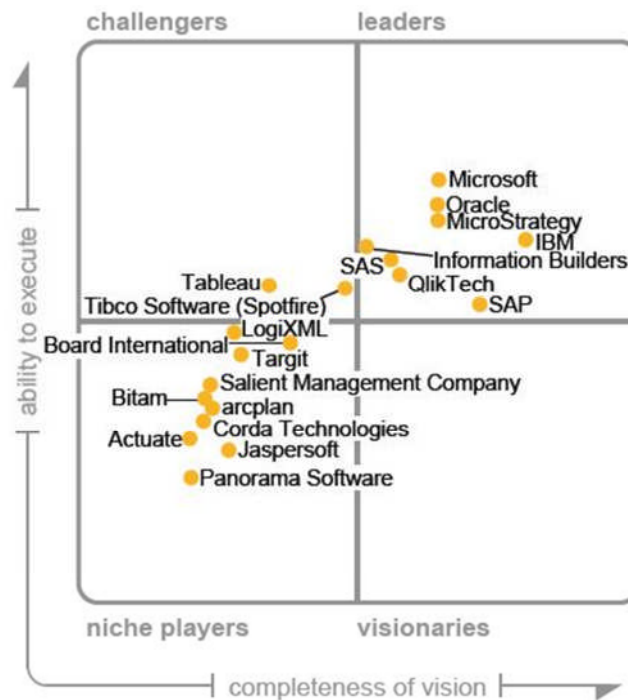


Figura 15 - Quadrante mágico para plataformas de BI (Sallam *et al.* 2011)

Para forma de exposição, decidiu-se explorar neste documento os fornecedores mais importantes actualmente no mercado de plataformas de BI, nomeadamente os que se inserem no quadrante dos *leaders* e que apresentam a maior habilidade para executar (Microsoft) e abrangência de visão (IBM).

2.4.1 Microsoft

A Microsoft apresenta-se como a organização que dentro dos líderes, possui a maior capacidade de “entregar à empresa implementações de larga escala que suportam uma ampla estratégia de BI (...) viabilidade e capacidade operacional capazes de serem fornecidas num âmbito global.” (Sallam *et al.* 2011). Tendo-se introduzindo no mercado de BI de uma forma mais séria em 2000, apresenta uma solução integrada aos seus clientes que atraí não só pela componente tecnológica bem estruturada e completa, como pelo preço competitivo apresentado (Sallam *et al.* 2011). Esta solução integrada assenta na conjugação do Microsoft Office, Microsoft SQL *Server* e Microsoft *Sharepoint*, como se consegue constatar na seguinte figura:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

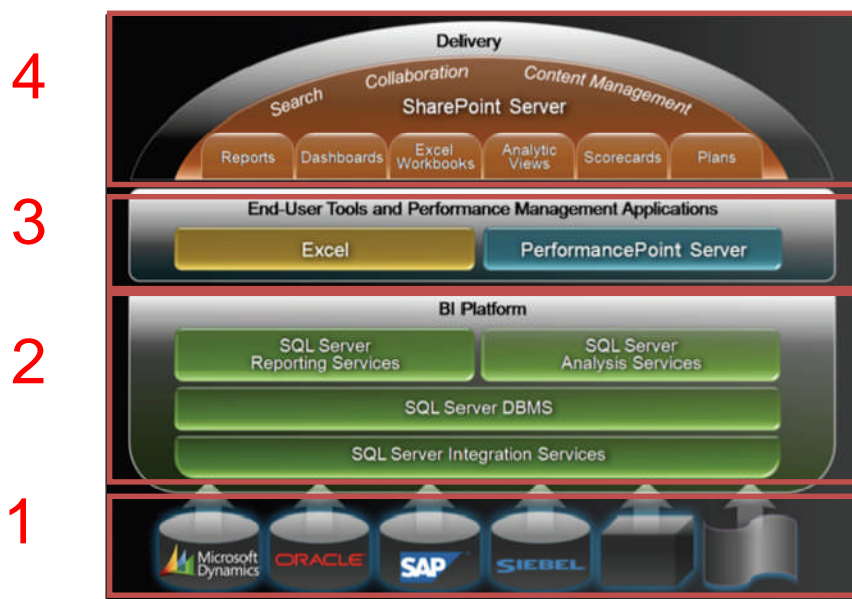


Figura 16 - Plataforma integrada de BI da Microsoft (Cardoso 2009)², (Lukawiecki 2009)]

Consegue-se então constatar quatro camadas distintas, sendo a primeira (1) as fontes utilizadas como entrada para a plataforma de BI, que cuida do ETL. A extração propriamente dita está ilustrada nas setas que se deslocam de (1) para (2), sendo a transformação e integração dos dados efectuada pela componente de *SQL Server Integration Services* (SSIS). A parte de armazenamento dos dados estará a cargo do gestor de base de dados *SQL Server DBMS*, que servirá de repositório e, que posteriormente, alimentará o *SQL Server Analysis Services* (SSAS), cuja função é a de criar os modelos dimensionais e respectivos cubos que servirão como *input* para o *SQL Server Reporting Services* (SSRS). O SSRS não suporta apenas dados modelados dimensionalmente mas também modelos relacionais. Toda esta relação de componentes associadas à plataforma de BI encontra-se ilustrada na seguinte figura:



Figura 17 - Plataforma de BI Microsoft (Cardoso 2009)²

Na camada (3) da figura 16, encontra-se a camada que serve de ferramenta para os utilizadores finais conseguirem ajustar da forma que mais lhes convém a informação que lhes chega a partir do nível (4) da figura 16, em forma de relatórios, *dashboards*, etc.

2.4.2 IBM

A IBM vem apresentar-se desta forma no quadrante mágico da Gartner como líder em termos de visionários. Tem a si associada uma vasta solução integrada de “(...) análise de desempenho e activos de *software* de gestão” (Sallam *et al.* 2011). Tendo lançado em Outubro de 2010 o IBM Cognos v10.1, este assenta “(...) na plataforma integrada do IBM Cognos 8 v4 que acrescenta um espaço de trabalho único para todos os conteúdos de BI denominado como *Business Insight*” (Sallam *et al.* 2011), sendo também o desempenho um aspecto que foi melhorado nesta nova *release* (Sallam *et al.* 2011). O IBM Cognos BI consegue-se representar no seguinte diagrama:

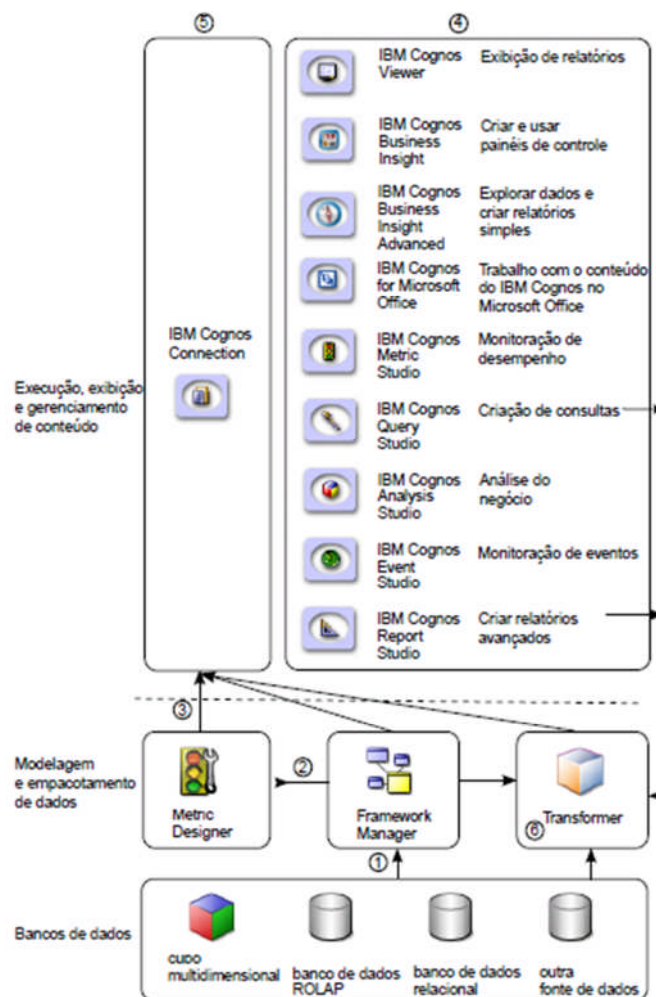


Figura 18 - Componentes do IBM Cognos BI (IBM 2010)

Como se pode observar, tem-se primeiro a camada de bancos de dados, que servirá como entrada para o *IBM Cognos Framework Manager*. “No *IBM Cognos Framework Manager*, o modelador

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

garante que os metadados sejam apresentados de um modo que os utilizadores possam entender. Os modeladores importam os metadados de um ou mais bancos de dados, e adicionam o modelo para corresponder às exigências do utilizador”. Ainda na parte de modelção e gestão dos dados tem-se o *Metric Designer* e o *Transformer*. Quando se passa para o *Metric Designer*, “o modelador identifica os itens de dados e cálculos que formam as metas e resultados reais para cada métrica”. Através do *Transformer*, fica “(...) possível nivelar consultas em pacotes publicados, consultas de relatórios do IBM Cognos BI e fontes de dados pessoais como o *software* de planilhas Microsoft Excel, para criar um modelo exclusivo e determinado de seus negócios”. Depois da parte de modelação ficar concluída, essa modelação é partilhada através do *IBM Cognos Connection*, para servir de *input* para todos os componentes inseridos no rectângulo 4 (IBM 2010).

2.5 Caso de estudo Janseen-Cilag

Neste último ponto do estado da arte, vai-se apresentar um caso de estudo na área realizado no passado. Este caso de estudo apresentado por Oliveira *et al.* (2006), tem como alvo a empresa Janseen-Cilag, que faz parte do grupo Johnson & Johnson. É uma empresa que produz produtos de saúde e fornece um conjunto de serviços para o consumidor, indústria farmacêutica, etc. A Janseen-Cilag é também líder na área de investigação da indústria farmacêutica, com mais de 19 mil colaboradores em 50 países (por volta do ano 2003) (Oliveira *et al.* 2006).

Como problemas tinha-se a orçamentação e planeamento financeiro do negócio que era baseado em folhas de cálculo excel, e devido à dimensão da organização, vários problemas foram identificados:

- Alto risco de erros;
- Alta complexidade estrutural;
- Dificuldade na manutenção dos acessos;
- Complexidade no carregamento dos dados para utilização em previsões e,;
- Dificuldade em gerir versões do mesmo exercício (ex: contabilístico) e comparação entre exercícios.

Tinham portanto, como objectivo principal, fornecer à organização uma ferramenta de planeamento eficaz e flexível, integrada com o sistema ERP actual, permitindo análises financeiras e um maior controlo nos seus actos de gestão. Como requisitos, foi identificado ser necessário utilizar uma ferramenta para carregar todos os dados, que fosse *web enabled*; poder

ser possível trabalhar em modo *offline* e,; capaz de suportar várias versões. Também foi identificado o desenvolvimento de um modelo orientado para a estrutura da organização.

Com isto, os resultados esperados centraram-se na geração automática de relatórios internos e corporativos, uma constituição de demonstrações financeiras por companhia e produto e, cubos multidimensionais que permitissem a análise e extração dos dados para as ferramentas de suporte utilizadas na empresa. A arquitectura da solução encontra-se ilustrada na seguinte figura:

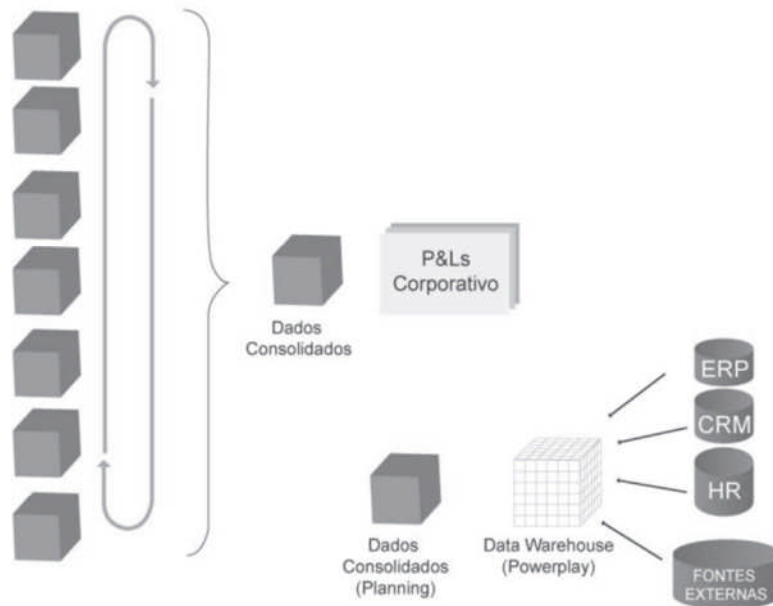


Figura 19 - Arquitectura da solução de BI da Janseen-Cilag

Com este projecto, a Janseen-Cilag encontrou como benefícios uma ferramenta única para o seu planeamento, que suporta várias versões, um menor tempo na formulação dessas versões e, um *deployment* a um nível global ao longo de toda a empresa.

Outros pontos também foram retirados deste trabalho, como a importância de se definir sempre o âmbito do projecto, envolver toda a organização em projectos desta natureza, fazer uma avaliação dos dados a trabalhar (validando-os) e investir o tempo necessário na fase inicial, compreendendo o negócio e a empresa em causa.

3 Abordagem metodológica

3.1 Ciclo de vida de um projecto de sistemas de apoio à decisão

Para o projecto adoptado na BI4ALL foi adoptada a metodologia desenvolvida por Kimball *et al.* (2008). A escolha por esta metodologia recaiu sobre o facto desta ter sido apresentada ao longo do Mestrado de Sistemas de Informação e ter sido igualmente seguida com maior profundidade numa das cadeiras de BI. Para além de ter tido esta influência no mestrado, foi possível constatar ao longo desta obra, que esta metodologia é também adoptada pelas organizações no desenvolvimento das suas soluções de BI. Por esta aceitação por parte do mundo académico e empresarial e, a sua evolução ao longo dos anos, fez todo o sentido em se adoptar e colocar em prática esta metodologia. “O ciclo de vida do Kimball fornece a estrutura para organizar as tarefas necessárias para implementar com sucesso um sistema DW/BI. Tal ciclo tem evoluído através de anos de experiência prática e baseia-se firmemente nas realidades que enfrentamos hoje” (Kimball *et al.* 2008).

Para Kimball *et al.* (2008), um projecto de DW/BI é necessário ter presente três vectores essenciais:

- Focar no negócio;
- Estruturar dimensionalmente os dados que são entregues ao negócio via relatório ou consultas *ad-hoc*;
- Desenvolver iterativamente todo o ambiente do DW de uma forma incremental, do que tentar realizar logo tudo de uma só vez;

Para isto, Kimball *et al.* (2008) desenvolveram um conjunto de iniciativas, ilustradas abaixo na seguinte figura:

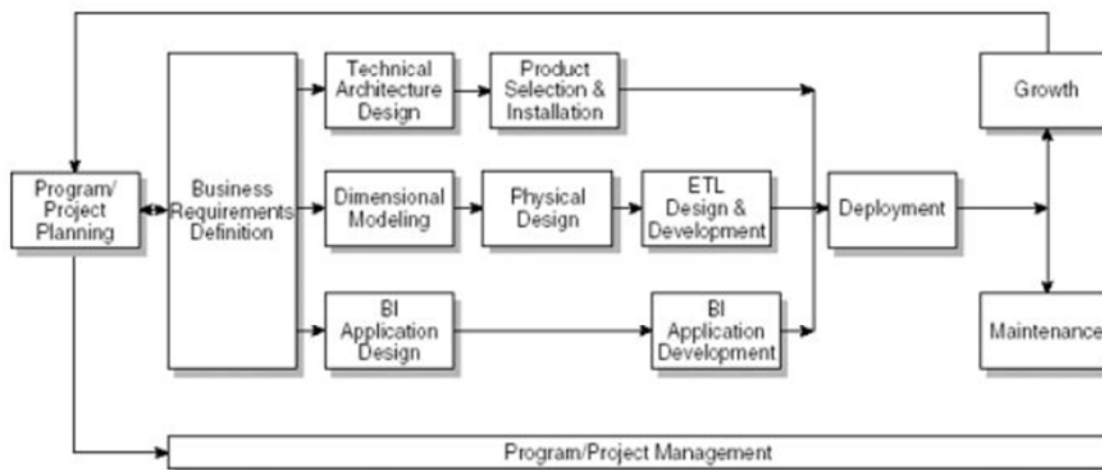


Figura 20 - Diagrama do ciclo de vida de Kimball *et al.* (2008)

Este diagrama permite que ao desenvolver um projecto de DW/BI, se torne possível identificar um conjunto de metas a atingir ao longo de todo o desenvolvimento, garantindo-se assim uma gestão do programa/projecto de uma forma controlada e eficaz “O ciclo de vida retrata a sequência de tarefas de alto nível necessárias para um design, desenvolvimento e implantação correctas de um sistema DW/BI. O diagrama mostra o roteiro geral, enquanto que cada caixa serve como objectivo a completar (...).” (Kimball *et al.* 2008).

3.1.1 Program/Project Planning

Este é o primeiro ponto de partida para qualquer desenvolvimento de um DW. É importante referir que um programa engloba um conjunto de projectos/iterações.

Como referido por Kimball *et al.* (2008), é difícil afirmar se primeiro vem o projecto ou o programa. “Em algumas organizações, existe um acordo para lançar um programa de DW/BI (...) priorizando qual o primeiro projecto a ser iniciado. Noutras situações, o financiamento é fornecido para um único projecto ou dois e, então, a necessidade de coordenação do programa é realizada posteriormente”.

Assuntos como se a organização estará pronta para um projecto de DW/BI – motivação, viabilidade, qualidade de dados, fraco apoio/envolvimento da gestão de topo, etc, são algo a apurar neste ponto da metodologia.

Segundo Kimball *et al.* (2008), é importante ter em conta três factores chave nesta avaliação:

1. Patrocinador(es) forte(s) proveniente(s) do negócio;
2. Uma motivação convincente por parte do negócio e,;
3. Viabilidade.

É igualmente importante reter que estes factores são muito pouco flexíveis, o que reforça a necessidade de validar estes aspectos nesta fase do projecto, pois é muito pouco provável que um projecto pouco viável no início, se vá depois tornar um projecto viável no futuro.

No primeiro factor chave acima citado, temos que um patrocinador da parte do negócio terá que possuir várias características que vão ser essenciais na sua classificação final. Entre essas características, destacam-se, segundo Kimball *et al.* (2008), as seguintes:

- 1) Ter a visão do potencial impacto de uma solução DW/BI;
- 2) Ter demonstrado um vasto conjunto de sucesso em vários projectos, sendo reconhecido por isso pelos restantes colaboradores;
- 3) É exigente, mas realista e colaborador em termos de suporte necessário;
- 4) Capaz de aceitar problemas a curto prazo, pois está focado no sucesso do projecto no longo prazo;
- 5) Capaz de se comprometer e,;
- 6) Capaz de fazer decisões difíceis e viver com as consequências.

No aspecto da viabilidade, segundo Kimball *et al.* (2008), viabilidade é usualmente entendida a um nível mais técnico, mas para este caso faz mais sentido associar a viabilidade orientado aos dados existentes. “Viabilidade usualmente refere-se exclusivamente à parte técnica. No entanto, a principal preocupação em relação à prontidão de uma iniciativa de DW/BI está relacionada com questões de viabilidade de dados”.

3.1.2 Program/Project Management

“A gestão de um programa/projecto garante que as actividades no ciclo de vida de Kimball permaneçam em desenvolvimento e em sincronia”. Tudo o que é relacionado com o controlo do âmbito de todo o projecto/programa é realizado neste ponto, assegurando que tudo é cumprido, o que não é cumprido e porquê, gerindo todos os incidentes ocorridos, evolução do projecto, etc, para ser sempre possível controlar o âmbito definido anteriormente, e em caso necessário, reajustá-lo.

Para Kimball *et al.* (2008), também é importante manter todos os envolvidos e afectados com todo o desenvolvimento actualizados sobre o progresso geral do programa/projecto, tentando controlar assim sempre as expectativas criadas na fase inicial. “Comunicação de uma forma contínua é fundamental para gerir as expectativas; as expectativas da parte de gestão são críticas para alcançar os objectivos do DW/BI”.

3.1.3 Business Requirements Definition

Este ponto é onde os utilizadores finais do sistema são questionados sobre as suas necessidades em termos de modelação, reunindo-se assim os requisitos necessários para o desenvolvimento futuro. “Sem esta compreensão, o DW/BI torna-se muitas vezes um exercício de futilidade para a equipa de projecto”.

É necessário perceber o negócio que está envolvido e perceber as suas necessidades em termos de modelação. Está-se então perante num dos pontos mais importantes de toda a metodologia.

3.1.4 Technical Architecture Design

“Criar uma arquitectura é um processo *top-down*. Primeiro percebe-se os requisitos de negócio, cria-se um modelo de alto nível que descreve a solução de uma perspectiva macro e, depois, é que se parte para os sub-sistemas (...) selecciona-se os produtos e depois implementa-se”.

Kimball *et al.* (2008), também referem que não se deve apenas focar na parte tecnológica como algo que está isolado e independente dos requisitos de negócio. “(...) os requisitos de negócio, ambiente técnico e, planeamento estratégico da direcção tecnológica a tomar - devem ser considerados simultaneamente para estabelecer o *design* da arquitectura técnica do DW/BI”.

Sobre a análise a nível de tecnologias a utilizar, isso também é realizado neste ponto, tal como o fluxo da informação deste o sistema fonte até ao utilizador final.

3.1.5 Product Selection and Installation

Aqui o objectivo centra-se em seleccionar todos os componentes desde *hardware*, até às ferramentas de extração de dados do DW. “(...) Componentes específicos da arquitectura como a plataforma de *hardware*, o sistema de gestão de base de dados, a ferramenta de ETL, ou as ferramentas de acesso e de *reporting* sobre os dados devem ser avaliados e seleccionados (...) depois devem ser instalados e testados”.

3.1.6 Dimensional Modeling

Depois de todo o levantamento realizado junto dos utilizadores chave de negócio, identifica-se as áreas de negócio da organização e as dimensões em causa. O resultado será um DW *Bus Matrix* que servirá como um “(...) diagrama da arquitectura de dados para assegurar que o DW/BI pode ser integrado e estendido ao longo do tempo por toda a organização”.

Depois desta visão mais geral sobre toda a organização, parte-se para uma análise mais centrada ao nível de cada área de negócio, analisando o nível de detalhe pretendido em termos de

Capítulo

3 Abordagem metodológica

granularidade e dimensionalidade desejada. Para isto, tem-se os quatro passos de Kimball *et al.* (2008) que se desdobram da seguinte forma:

Passo 1 – Escolher o processo de negócio;

- Dos processos de negócio identificados no *DW Bus matrix*, começa-se por escolher um. “(...) escolher o processo a modelar. Esta escolha já deverá ter sido realizada pelo negócio na fase de priorização de actividades (...).

Passo 2 – Definir o grão;

- Para o processo ou sub-processos seleccionados no passo 1, declara-se o grão aí presente.

Passo 3 – Identificar as dimensões;

- Neste passo identificam-se as dimensões que vão de encontro com o grão definido no passo 2.

Passo 4 – Identificar os factos.

- Finalmente, no passo final, procede-se à identificação dos factos também especificados no grão no passo 2.

3.1.7 Physical Design

Neste ponto são definidas as “estruturas físicas, incluindo a preparação para o ambiente de base de dados e instanciando a segurança apropriada”.

Tópicos como a forma como as chaves estrangeiras e chaves primárias são geridas também é aqui definido, convenções de nomenclatura das tabelas e colunas seguirem os *standards* adoptados, modelos em estrela *versus* modelos floco-de-neve, etc.

3.1.8 ETL Design and Development

Neste ponto tem-se como objectivo realizar a fase de ETL, onde 70% do risco de projectos de DW/BI se encontram alojados.

3.1.9 BI Application Design

Aqui pretende-se identificar através dos utilizadores finais, a forma como pretendem navegar pelos dados na aplicação de BI e como pretendem ter os dados disponibilizados/dispostos nessa mesma aplicação. “Aplicações de BI são o veículo que entrega o valor do negócio contido na solução de DW/BI, ao invés de apenas entregar os dados”.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

3.1.10 *BI Application Development*

“(…) configurar os metadados do negócio e a infra-estrutura das ferramentas e, em seguida, construir e validar as aplicações de BI”.

3.1.11 *Maintenance*

Aqui o que se encontra em produção é tido em conta, tendo sempre a preocupação de ter o actual sistema a funcionar correctamente, “quer em termos de desempenho quer em divulgação junto dos utilizadores finais e suporte aos mesmos (...) para manter o sistema com um desempenho bom (...) melhoramento do desempenho, manutenção dos índices e, *backup* do sistema (...) focando-se nos utilizadores de negócio com suporte permanente”.

3.1.12 *Growth*

Este é o ponto onde novas necessidades são identificadas e onde o actual sistema necessita de ser incrementado em termos de valor para o negócio. A partir dessa identificação, volta-se ao ponto associado ao planeamento do programa/projecto, recomeçando assim um novo ciclo, ou neste caso, projecto.

4 Caso de estudo BI4ALL

Neste capítulo vai ser exposto os resultados obtidos através da aplicação directa da metodologia explorada no capítulo anterior (3) deste documento. Em certos casos, será dividido em sub-tópicos quando se explorar de uma forma paralela as duas áreas funcionais que estão a ser alvo de estudo: Controlo e redução de custos e, gestão de projectos.

Importante referir apenas alguma documentação definida para cada iniciativa recomendada por Kimball *et al.* (2008) na sua metodologia foram desenvolvidos. Isto deveu-se ao facto de apenas alguma fazer sentido ser desenvolvida para este trabalho, não se fazendo da metodologia algo como um “*must step-by-step*”, mas sim utilizando o que é mais útil neste caso de estudo ao longo de todas as iniciativas. Essa documentação encontra-se presente ao longo deste documento.

4.1 Program/Project management

“O risco é reduzido em primeiro avaliar se a sua empresa está pronta para uma iniciativa de DW/BI” (Kimball *et al.* 2008).

4.1.1 Avaliar a disponibilidade da organização para um projecto de DW/BI

Em qualquer projecto de DW/BI, é necessário avaliar a disponibilidade da organização para um projecto destas características. Segundo Kimball *et al.* (2008), é bastante importante realizar esta avaliação pois em caso negativo, evita-se investimentos realizados de uma forma precoce e, que não irá trazer uma mais valia para a organização, existindo assim um alocação de recursos desnecessários.

No caso da BI4ALL, foram identificados como patrocinadores o Dr. Andro Moreira, o Dr. José Oliveira e o Eng^o Rui Afeiteira. Sendo a empresa em causa uma empresa que presta serviços de BI, *corporate governance* e *risk management* há já algum tempo e, a experiência que estes patrocinadores já possuem no vasto conjunto de projectos em que já se viram envolvidos, as características descritas de 1) a 7) no ponto 3.1.1 estão completamente cobertas e validadas positivamente, concluindo-se serem patrocinadores fortes e líderes na sua função interna na organização.

O segundo factor chave centra-se então na motivação por parte do negócio, em que, segundo Kimball *et al.* (2008), “um DW é um possibilitador para endereçar requisitos críticos do negócio”.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Segundo o que foi possível apurar, a organização apresenta várias motivações para se avançar com este projecto, as quais já foram apresentadas e descritas nos pontos 1.2 Problema e 1.3 Objectivos. Conclui-se que, no que se trata de motivação por parte da organização, essa encontra-se plenamente identificada e justificada.

Sobre a viabilidade, foi possível apurar que os dados estão descentralizados, estando uns localizados em folhas de cálculo excel, outros assentes no ERP ArtSoft, outros assentes em Microsoft Access e por fim, outros alojados em SQL Server (*Active Directory*). A complexidade associada a estes dados foi validada pelos patrocinadores do projecto, tendo-se concluído que existia um problema relativo à qualidade dos dados a ser trabalhados e, que isto poderia dificultar a componente prática.

De uma maneira geral, é possível assegurar que a cultura existente na organização é orientada para decisões baseadas em factos, tendo sido promovido a intenção de se partir para este projecto de forma a colmatar o sub-aproveitamento da informação existente na organização.

4.1.2 Endereçar deficiências e determinar os próximos passos

“(…) é preciso avaliar a prontidão da organização. Estratégias para redução de riscos e planos de contingência devem ser desenvolvidos para tratar qualquer deficiência detectada” (Kimball *et al.* 2008).

Como factor mais importante e crítico tem-se o forte apoio do(s) patrocinador(es) do negócio. É até sugerido, que mesmo que se parta para um projecto de DW/BI que tente satisfazer um vasto conjunto de necessidades, isso nunca irá ser superior se fosse com o apoio do tal patrocinador: “Até a solução mais elegantemente desenhada de DW/BI não consegue sobrepor a falta de apoio por parte do negócio”.

Sendo assim, Kimball *et al.* (2008) descrevem um conjunto comum de cenários que permite à organização avaliar se avança com o projecto e, que acções poderá tomar sobre as deficiências encontradas. Esses cenários vão ser utilizados para este caso de estudo:

1. “Motivação do negócio convincente”

“Do nosso ponto de vista, os sistemas DW/BI são bem sucedidos se forem aceites pelo negócio para ajudar no seu suporte à decisão”.

É possível constatar que este cenário está presente neste caso de estudo, o que valida a motivação para o desenvolvimento deste caso de estudo para os próximos passos.

2. “Pobre qualidade de dados”

“Mesmo com um patrocinador forte e uma necessidade para o negócio que se mostre atraente, sem dados com qualidade, o projecto torna-se em missão impossível”.

Para Kimball *et al.* (2008), se neste ponto for verificado a existência de uma qualidade de dados pobre, o ponto 1. fica altamente comprometido.

Pelo o que se apurou neste caso de estudo, não parece existir este constrangimento ao nível dos dados dos sistemas fonte. Se para algum requisito que for identificado nas entrevistas tiver dados a sí associados que possam comprometer o sucesso final, deve-se passar por identificar outro requisito como sendo um novo requisito core a ser implementado: “Ao invés de se prosseguir com este tipo de projecto, seria mais sensato identificar outro requisito de alto valor para o patrocinador com menor obstáculos em termos de viabilidade dos dados”. Para as áreas seleccionadas para desenvolvimento, apenas alguns problemas foram detectados relativamente a qualidade de dados (identificado no ponto 4.1.1) mas que não era crítico e não comprometia o sucesso do projecto.

3. “Silos de dados antigos, isolados e de pouco desempenho”

“Muitas organizações já têm construídos um DW ou DMs”.

Na organização já existe um DW e um DM (Gestão de projectos) mas que se encontra desactualizado face às actuais necessidades levantandas.

4. “Patrocinador forte, Strong sponsor, necessidade do negócio convincente, e qualidade dos dados”

“Parabéns! Tens todos os factores chave para uma iniciativa deste tipo validadas.”.

É possível constatar que este cenário está presente neste caso de estudo, pois nos três pontos anteriores teve-se uma avaliação positiva por parte da organização.

4.2 Business Requirements Definition

“Utilizadores do negócio e seus requisitos têm impacto em praticamente todas as decisões tomadas na fase de design e implementação (...) eles são tão importantes para um *data warehousing* bem sucedido” (Kimball *et al.* 2008).

4.2.1 Entrevistas

Para este ponto do trabalho, recorreu-se à realização de um questionário que teve por base da sua construção muitos dos tópicos recomendados por Kimball *et al.* (2008). Antes do desenvolvimento do questionário, os utilizadores alvo para o efeito do levantamento de requisitos foram identificados, sendo eles os maiores beneficiados com todo o desenvolvimento que viria a ser realizado. Esses utilizadores foram o Dr. José Oliveira – CEO da BI4ALL, Dra. Dora Palma – Assessora de Comunicação, Dr. Andro Moreira – *Partner* da BI4ALL, Dr. Rui Gorgueira – *Manager* da BI4ALL e o Eng^o Rui Afeiteira – Consultor da BI4ALL.

O questionário usado nas entrevistas agrupou-se nas seguintes categorias:

1. Prioritização dos processos de negócio;
 - Identificação dos modelos de negócio chave e os seus sistemas fonte.
2. Objectivos de negócio e problemas;
 - Métricas do negócio e oportunidades.
3. Análise de requisitos.
 - Relatórios-chave & tipos de análises e, limitações actuais.

4.2.2 Resultados obtidos das entrevistas

Após as entrevistas conseguiu-se constatar que os duas áreas funcionais a modelar dimensionalmente para este trabalho seriam as de controlo e redução de custos (componente de chamadas) e gestão de projectos.

▪ Levantamento da área de controlo e redução de custos

Sobre a área de controlo e redução de custos foi possível constatar que nada existe desenvolvido na organização em termos de modelação dimensional, sendo até hoje este controlo inexistente devido ao árduo trabalho manual associado em ter toda a informação necessária em tempo útil e em formatos trabalháveis. Por esta ser uma área que reúne falhas em termos de disponibilização de informação, foi decidido dar-lhe maior prioridade.

As dimensões de análise identificadas para este modelo foram as seguintes:

- Tempo;
- Colaborador;
- Período;
- Tipo de chamada;
- Tipo de comunicação.

Capítulo 4 Caso de estudo BI4ALL

As métricas elementares que foram levantadas neste modelo são as seguintes:

- Custo da chamada;
- Custo da chamada real;
- Duração em minutos;
- Volume de dados e,;
- Número de ocorrências.

Para mais informação associada a estas métricas e dimensões, por favor verificar o **Anexo A – Hierarquias e o Anexo B – Metadata (DW)**.

Os indicadores levantados estão descritos na tabela abaixo:

| Indicador | Fórmula de Cálculo | Unidade | Aditiva | Periodicidade de Cálculo | Valor Base |
|---|---|---------|---------|--------------------------|---|
| Valor a pagar de chamadas nacionais | $\sum(\text{Duração em minutos})$ para os tipos de chamada iguais a Nacionais * 0.03 | Euro | S | Mensal | Duração em minutos |
| Valor a pagar de chamadas internacionais | $\sum(\text{Custo da chamada})$ para os tipos de chamada iguais a Internacionais | Euro | S | Mensal | Custo da chamada |
| Valor a pagar de chamadas recebidas em roaming | $\sum(\text{Duração em minutos})$ para os tipos de chamada iguais a Chamadas Recebidas * 0.142 | Euro | S | Mensal | Duração em minutos |
| Valor a pagar de chamadas efectuadas em roaming | $\sum(\text{Duração em minutos})$ para os tipos de chamada iguais a Chamadas Efectuadas * 0.36 | Euro | S | Mensal | Duração em minutos |
| Valor a pagar SMS | $\sum(\text{Número de ocorrências})$ para os tipos de chamada iguais a Mensagem Escrita Enviada – COD_PLAFOND_SMS * 0.108 | Euro | S | Mensal | Numero de ocorrências |
| Valor a pagar MMS | $\sum(\text{Número de ocorrências})$ para os tipos de chamada iguais a Mensagem Multimédia Enviada * 0.03 | Euro | S | Mensal | Numero de ocorrências |
| Valor a pagar Roaming dados | $\sum(\text{Número de ocorrências})$ para os tipos de chamada iguais a Dados Roaming Zona ou Dados Roaming Zona 1 * 0.00607 | Euro | S | Mensal | Numero de ocorrências |
| Valor total a pagar das chamadas | $\sum(\text{Custo da chamada})$ | Euro | S | Mensal | Custo da chamada |
| Valor total a pagar | Valor total a pagar das chamadas + Valor a pagar SMS + Valor a pagar MMS + Valor a pagar Roaming dados | Euro | S | Mensal | ND (baseia-se nos indicadores pré-calculados) |

Tabela 3 - Resumo dos indicadores do modelo de controlo e redução de custos

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Para este modelo, foram identificados três tipos distintos de fontes de informação, independentes entre si. Essas fontes são as seguintes:

1. *Active Directory*;
 - a) Assente em *SQL Server 2005*.
2. Ficheiro Optimus (Formato Excel);
 - a) Este ficheiro é enviado mensalmente pela Optimus para a BI4ALL, contendo toda a informação associada às chamadas realizadas pelos seus colaboradores.
3. Ficheiro *Plafonds* de cada utilizador (Formato Excel).
 - a) Este ficheiro é mantido manualmente pela BI4ALL onde estão definidos os *plafonds* associados a cada colaborador.

A seguinte ilustração exemplifica como é que as fontes de informação estão dispostas e quais as tabelas/*sheets* necessárias para o modelo:

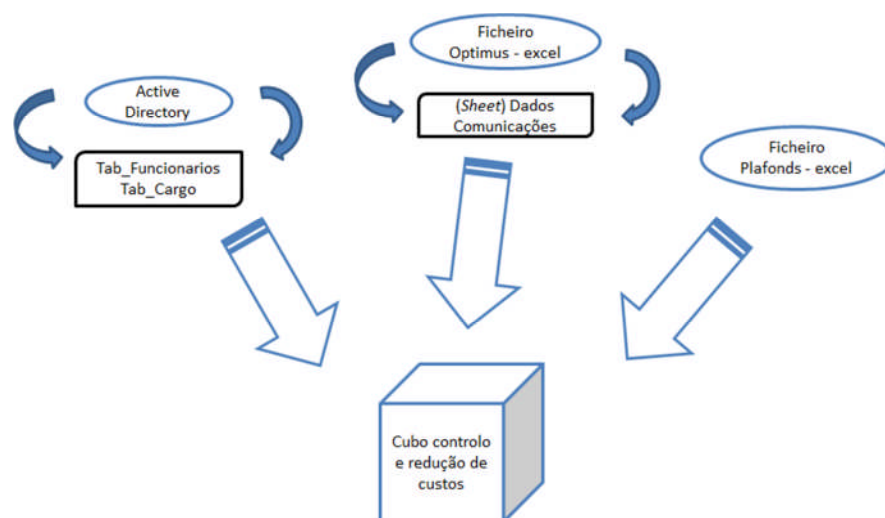


Figura 21 - Diagrama das fontes de informação do modelo de controlo e redução de custos

▪ Levantamento da área de gestão de projectos

Relativamente a esta área funcional, já existia um modelo desenvolvido mas que se encontrava desactualizado e necessitava de re-estruturação tanto a nível de informação disponibilizada, como a nível de desenvolvimento prático, onde a nomenclatura utilizada na altura não foi a mais correcta, tornando a gestão da mudança no que existia implementado mais difícil para os consultores que desenvolvem este tipo de sistemas. Foi então pedido pela BI4ALL refazer todo este modelo, começando do zero. Começou-se por avaliar se a informação que actualmente era modelada estava ainda correcta, e foi também efectuado o levantamento sobre o que faltava implementar de forma a colmatar a lacuna existente.

As dimensões de análise identificadas para este modelo foram as seguintes:

- Tempo;
- Colaborador;
- Projecto;
- Cliente;
- Tipo de hora.

As métricas elementares que foram levantadas neste modelo são as seguintes:

- Número de horas;
- Número de horas suplementares;
- Número de dias consumidos;
- Número de horas facturáveis;
- *Target*.

Para mais informação associada a estas métricas e dimensões, por favor verificar o **Anexo A – Hierarquias e o Anexo B – Metadata (DW)**.

Os indicadores levantados estão descritos na tabela abaixo:

| Indicador | Fórmula de Cálculo | Unidade | Aditiva | Periodicidade de Cálculo | Valor Base |
|------------------------------|---|---------|---------|--------------------------|------------|
| Capacidade máxima instalada | Número de consultores * número de dias úteis do mês * 8 horas | # | S | Mensal | NA. |
| Capacidade disponível | Número de dias trabalháveis por consultor * 8 horas | # | S | Mensal | NA. |
| Capacidade utilizada/ocupada | Número de horas úteis (horas facturáveis + prospects) | # | S | Mensal | NA. |
| Capacidade facturada | Número de horas facturadas (horas facturáveis) | # | S | Mensal | NA. |
| Taxa de disponibilidade | Capacidade disponível / Capacidade máxima instalada | % | N | Mensal | NA. |
| Taxa de ocupação | (Capacidade utilizada/ocupada) / Capacidade disponível | % | N | Mensal | NA. |
| Taxa de facturação | Capacidade facturada / Capacidade disponível | % | N | Mensal | NA. |

Tabela 4- Resumo dos indicadores do modelo de gestão de projectos

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Para este modelo, a fonte de informação baseia-se num ERP denominado ArtSoft, cuja tecnologia de gestão de base de dados é o Microsoft *Access*. Os dados aí contidos apresentam boa qualidade, sendo que a informação se encontra localizada em duas bases de dados separadas: dbSgo.mdb e dbSgo_CP.mdb. As tabelas necessárias para a extracção dos dados encontram-se ilustradas na seguinte figura:

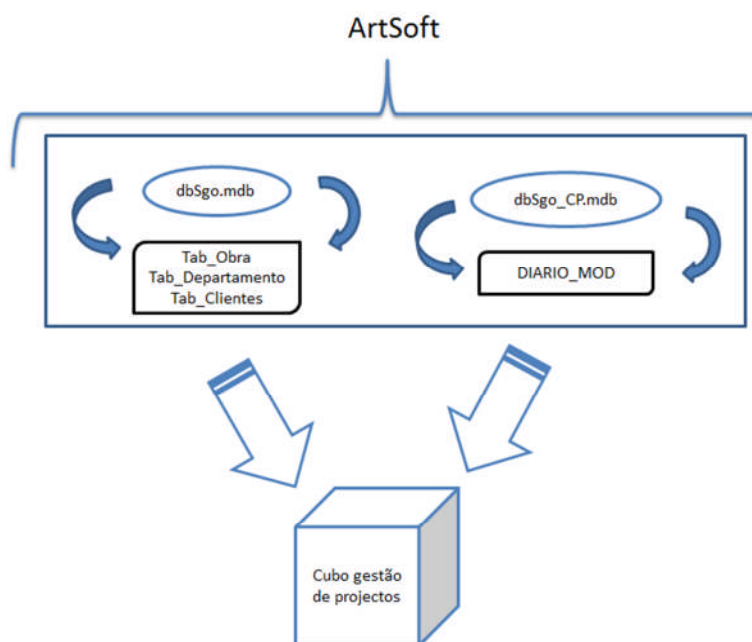


Figura 22 - Diagrama das fontes de informação do modelo de gestão de projectos

4.2.3 Enterprise DW Matrix

Para Kimball *et al.* (2008), conseguir demonstrar o que é mais importante atingir na hora de se arrancar para um projecto de DW/BI é crucial. “Se a construção do DW de uma só vez é assustador e contruí-lo como peças isoladas derrota o objectivo geral, o que deve ser feito?”. É então sugerido a utilização da *Enterprise DW Bus matrix*, que permite exemplificar de uma forma muito simples e directa, a verdadeira dimensão e impacto que o projecto de DW/BI irá ocupar.

Para este caso de estudo, a matriz desenvolvida encontra-se seguinte tabela:

| Áreas de negócio | Dimensões | | | | | | | | |
|-------------------------------------|-----------|-------------|---------|-----------------|---------------------|----------|-------------|---------|-------------------|
| | Data | Colaborador | Periodo | Tipo de chamada | Tipo de comunicação | Projecto | Tipo de hor | Cliente | Local de trabalho |
| Controlo e redução de custos | x | x | x | x | x | | | | |
| Chamadas Optimus | x | x | x | x | x | | | | |
| Gestão de projectos | x | x | | | | x | x | x | x |
| Actividades | x | x | | | | x | x | x | x |
| Actividades facturáveis | x | x | | | | | | | |
| Utilização de mão de obra | x | | | | | x | | | |
| Targets | | | | | | x | | | |

Tabela 5 – Enterprise DW Bus Matrix

Depois de analisada a matriz, foi decidido iniciar o desenvolvimento pelo modelo de controlo e redução de custos – Chamadas Optimus. Esta decisão foi tomada com base na complexidade deste modelo (que contém apenas um sub-modelo) e por ser algo que ainda não se encontra modelado, os utilizadores finais referiram a urgência em se ter toda a informação da componente de chamadas modelada para a primeira fase de desenvolvimento.

4.3 Technical Architecture Design

Nesta fase, a informação ao nível de arquitectura foi algo que foi transmitido numa fase muito inicial do trabalho, sendo um aspecto positivo e antecipando necessidades ou surpresas que se poderia vir a ter no desenrolar do caso de estudo. Posto isto, a seguinte figura ilustra de uma forma macro, os componentes desde a fonte de dados até à camada de apresentação:

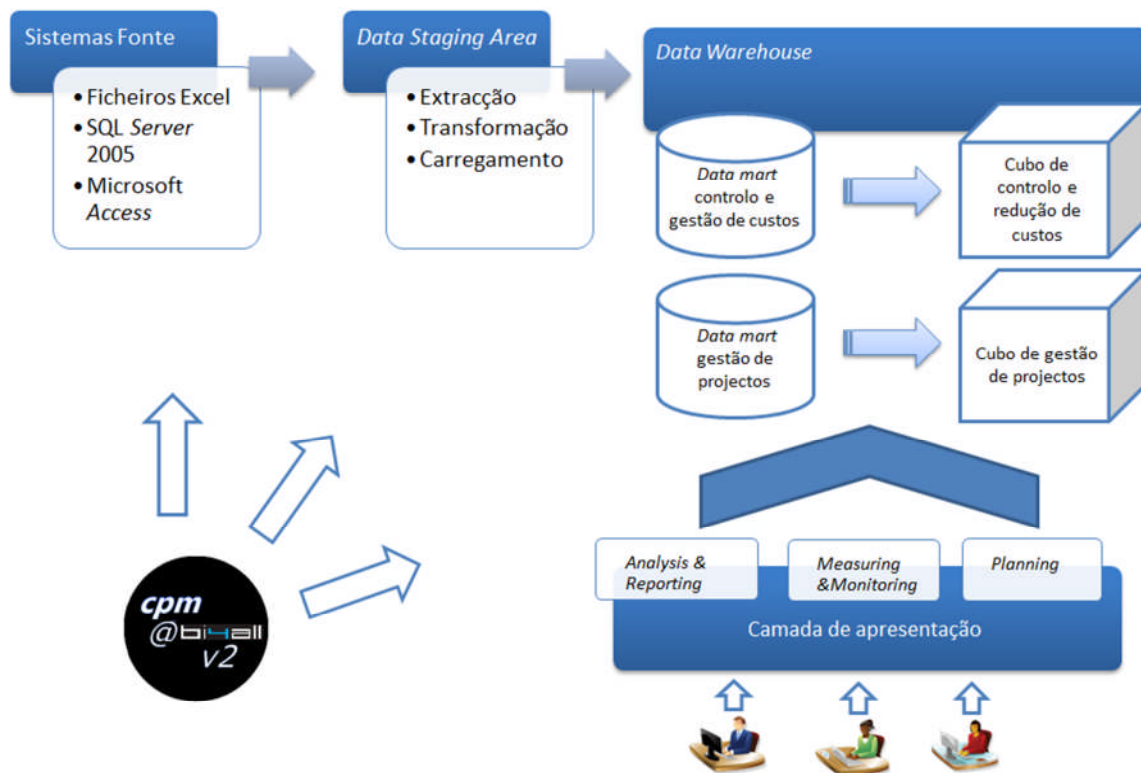


Figura 23 - Modelo macro da arquitectura desenvolvida

Através desta perspectiva macro, os sub-sistemas que se conseguem extrair da figura 23, encontram-se ilustrados na seguinte figura:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

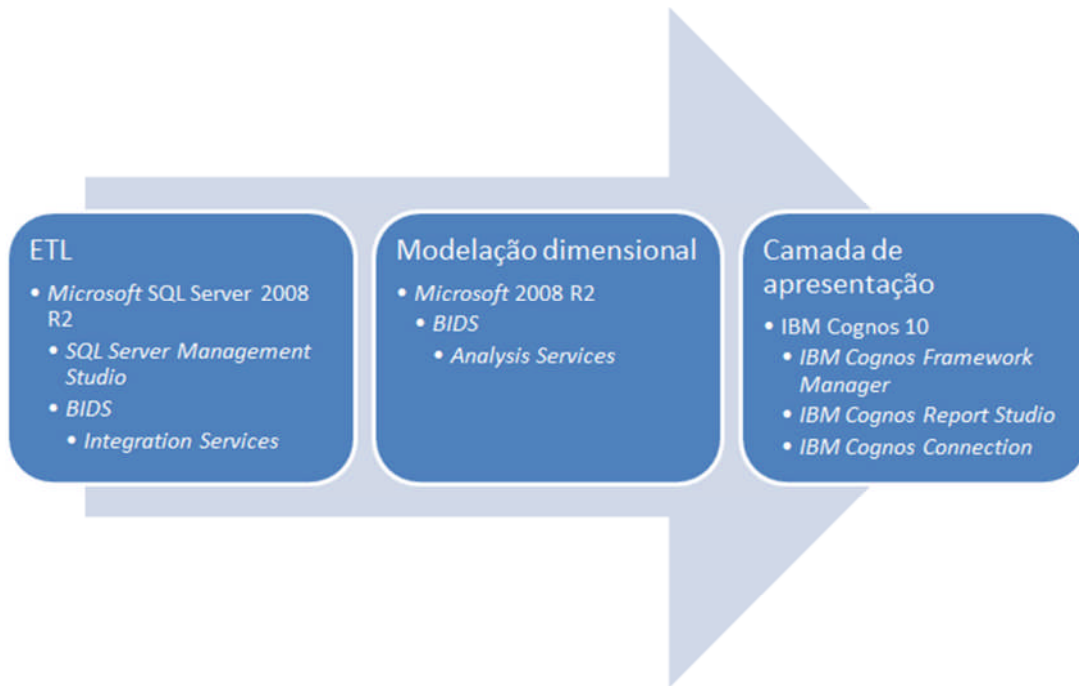


Figura 24 - Sub-sistemas da arquitectura do caso de estudo

Como se pode observar, está-se na presença de três sub-sistemas que compõem o modelo macro apresentado acima. No primeiro vai-se recorrer à tecnologia *Microsoft SQL Server 2008 R2* que fornece um gestor de bases de dados. A arquitectura aqui instanciada segue a metodologia adoptada pela BI4ALL nos seus projectos e, define-se da seguinte forma:

Base de dados de *Staging*

- Todas as tabelas temporárias para as dimensões e factuais encontram-se aqui criadas;
- É utilizada a nomenclatura SA_DIMENSAO e SA_FACTUAL.

Base de dados de *Config*

- Aqui são criadas tabelas de configuração dos *packages* desenvolvidos no SSIS, tornando o projecto mais flexível para a mudança – *SSIS Configurations*.
- Tabelas auxiliares são também aqui armazenadas (ex: AUX_DIM_TEMPO).

Base de dados de *DW*

- Aqui são armazenadas as tabelas que alimentam e definem o *datamart*. Todas as dimensões e factuais possuem a nomenclatura DIM_Dimensão e FCT_Factual.
- A camada de SSAS é alimentada através daqui pelas *views* que são criadas depois de carregadas as dimensões e as factuais.

Base de dados de *Log*

- No carregamento das factuais existe uma tabela de *log* por cada onde são armazenadas as linhas que não cruzaram com uma ou mais dimensões na fase de *loading*. Essas tabelas seguem a nomenclatura LOG_FCT_Factual.
- Também é aqui gerida uma tabela LogExecucao que armazena informação importante ao longo do ciclo de vida de processamento de um *package* SSIS e uma tabela de LOTE que armazena informação sobre qual o sistema fonte da extração, data de início e de fim de execução, etc.

Figura 25 - Estrutura de registo da informação ao longo do processo de ETL

Ainda dentro do subsistema ETL tem-se a componente de SSIS que irá conter todos os pacotes relativos ao ETL e que vão preencher com dados as bases de dados exploradas em pontos anteriores. A estruturação destes pacotes baseia-se numa camada dedicada à extração e transformação (ET), uma camada de carregamento (L - “*loading*”) para as dimensões e factuais no DW, e uma última camada com a gestão de processamento dos pacotes de ET+L. Para este caso de estudo, e seguindo esta arquitectura e nomenclatura, obteve-se a seguinte estrutura como *output*:

- Uma solução SSIS composta por três projectos:
 - BI4ALL@CPM-OPTIMUS_ET
 - BI4ALL@CPM-OPTIMUS_L
 - BI4ALL@CPM-OPTIMUS_MASTER

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

O BI4ALL@CPM-OPTIMUS_ET contém todos os pacotes que vão extrair os dados aos sistemas fonte, tratá-los e, posteriormente entregar o resultado final às tabelas temporárias contidas na *staging area*. A ilustração gráfica do BI4ALL@CPM-OPTIMUS_ET está indicada no **Anexo C - Pacotes de extração e transformação da informação dos sistemas fonte para a *staging area***.

O BI4ALL@CPM-OPTIMUS_L contém todos os pacotes que vão extrair os dados da *staging area* preenchida anteriormente pelo BI4ALL@CPM-OPTIMUS_ET. Ao extrair, para cada dimensão irá actualizar ou inserir novos registos, enquanto que as factuais vão ser inseridas novas linhas (ou apagadas e inseridas novamente), sendo essas linhas cruzadas com as dimensões – sendo então as chaves substitutas contidas nas dimensões passadas como chaves estrangeiras para as factuais. A ilustração gráfica do BI4ALL@CPM-OPTIMUS_L está indicada no **Anexo C - Pacotes de carregamento das dimensões e factuais**.

4.4 Dimensional Modeling

Depois de se ter reunido toda a informação necessária na fase de levantamento de requisitos, chega agora a altura de verificar tudo o que foi recolhido, validar, e proceder à modelação dimensional juntamente com todos os intervenientes no processo.

Seguindo a metodologia de Kimball *et al.* (2008), realizou-se para cada modelo os quatro passos de Kimball.

4.4.1 Os quatro passos de Kimball – Área de gestão e controlo de custos

Na metodologia de Kimball *et al.* (2008) é referido como primeiro passo identificar o processo de negócio. Neste caso, não se qualifica controlo e redução de custos ou gestão de projectos como processos de negócio, mas sim como áreas funcionais na organização. Por esta razão, são referidos como áreas funcionais ao invés de processos. Sendo assim, e colocando em prática para este caso de estudo a área de negócio a modelar primeiro no ponto 4.2, tem-se:

- **Passo 1 – Escolher a área funcional**

Controlo e gestão de custos – componente de chamadas.

- **Passo 2 – Declarar o grão**

A duração, custo (real e efectivo) e volume de uma chamada realizada num dado dia, num dado período do dia por um colaborador da organização que realizou um certo tipo de comunicação com um certo tipo de categorização de chamada efectuada.

▪ **Passo 3 – Identificar as dimensões**

Dimensão colaborador, período, tempo, tipo de chamada e tipo de comunicação.

▪ **Passo 4 – Identificar os factos**

Duração em minutos da chamada, o volume de dados gastos na chamada e o custo real e efectivo da chamada realizada.

4.4.2 Diagrama dimensional – Modelo de controlo e redução de custos

Após se efectuar os quatro passos de Kimball para a área de gestão e controlo de custos, procedeu-se à elaboração do modelo em estrela contendo as métricas e dimensões identificadas anteriormente. Este modelo é baseado numa tabela de factos *snapshot* periódica (com periodicidade mensal). O modelo encontra-se na seguinte figura:



Figura 26 - Diagrama dimensional do modelo de controlo e redução de custos

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Na construção deste modelo dimensional (e todos os outros) teve-se em atenção a utilização da nomenclatura utilizada pela BI4ALL nos seus projectos. Cada atributo da dimensão contém um código e a sua respectiva descrição. Muitas das vezes o código representa o valor que vem do sistema fonte, onde por vezes é necessário transformar/ajustar esse valor para algo mais perceptível. O campo descritivo representa o que é que os utilizadores finais interpretam que seja o tal código, ou seja, este campo é sempre transformado de forma a ir de encontro com o desejo os utilizadores. Verificar o **Anexo B – Metadata (DW)** para exemplos de preenchimento. Na definição da solução de exploração de dados, mais propriamente da dimensão no SSAS, o campo COD_NOME será a chave, sendo o campo a aparecer ao navegar no cubo o DSC_NOME. De referir que a DIM_TEMPO apenas contém alguns atributos do que é fisicamente guardado (foi uma opção tomada em termos de visualização neste documento). Nas factuais também se adoptou a nomenclatura de utilizar IND_X para métricas e INF_X para campos auxiliares (como *flags*).

4.4.3 Matriz de validação de métricas

Para a validação do modelo recorreu-se ao uso de uma matriz de validação de métricas, em que se cruza as métricas com as dimensões identificadas anteriormente. A matriz de validação para o modelo de controlo e redução de custos encontra-se na seguinte tabela:

| Controlo e redução de custos | Dimensões | | | | | | | | | | |
|---|-----------|------|-------------|---------|-----------------|---------------------|----------|---------|---------|--------------|----------|
| | Tipo | Data | Colaborador | Periodo | Tipo de chamada | Tipo de comunicação | Projecto | Cliente | Local d | Tipo de hora | Incluir? |
| Métricas elementares | # | x | x | x | x | x | | | | | Sim |
| Custo da chamada | # | x | x | x | x | x | | | | | Sim |
| Duração em minutos | # | x | x | x | x | x | | | | | Sim |
| Número de ocorrências | # | x | x | x | x | x | | | | | Sim |
| Volume de dados | # | x | x | x | x | x | | | | | Sim |
| Métricas agregadas | | | | | | | | | | | |
| Valor a pagar de chamadas efectuadas em roaming | Σ | x | x | x | x | x | | | | | Sim |
| Valor a pagar de chamadas internacionais | Σ | x | x | x | x | x | | | | | Sim |
| Valor a pagar de chamadas nacionais | Σ | x | x | x | x | x | | | | | Sim |
| Valor a pagar de chamadas recebidas em roaming | Σ | x | x | x | x | x | | | | | Sim |
| Valor a pagar MMS | Σ | x | x | x | x | x | | | | | Sim |
| Valor a pagar Roaming dados | Σ | x | x | x | x | x | | | | | Sim |
| Valor a pagar SMS | Σ | x | x | x | x | x | | | | | Sim |
| Valor a pagar das chamadas | Σ | x | x | x | x | x | | | | | Sim |
| Valor total a pagar | Σ | x | x | x | x | x | | | | | Sim |

Tabela 6 - Matriz de validação de métricas do modelo de controlo e redução de custos

Através da matriz apresentada acima, consegue-se então provar o modelo descrito anteriormente se encontra validado e que as métricas se encontram dentro do grão definido nos quatros passos de Kimball. De referir que a coluna “**Incluir?**” refere-se ao facto de a métrica estar incluída no cubo, ou de ser apenas calculada quando na consulta sobre os dados. Ou seja, se for incluída, será pré-calculada antes de se navegar no cubo (existe armazenamento físico), senão, é calculada em tempo real.

4.4.4 Os quatro passos de Kimball – Modelo de gestão de projectos

- **Área de gestão de projectos - Actividades**

Para a área de gestão de projectos temos quatro sub-áreas que estão apresentadas na *Enterprise DW Bus matrix* em 4.2.3. Sendo assim, vai-se efectuar o processo dos quatro passos de Kimball para cada uma das sub-áreas (Actividades, Actividades facturáveis, Utilização de mão de obra e *Targets*).

- **Passo 1 – Escolher a área funcional**

Gestão de projectos – Actividades

- **Passo 2 – Declarar o grão**

A duração em horas (facturáveis, normais e suplementares) e dias consumidos num dado projecto, por um certo colaborador, realizado num dado local de trabalho para um determinado cliente num determinado dia.

- **Passo 3 – Identificar as dimensões**

Dimensão colaborador, projecto, local de trabalho, cliente e tempo.

- **Passo 4 – Identificar os factos**

Duração em horas (normais, facturáveis e suplementares), número de dias consumidos.

Para esta sub-área de gestão de projectos e depois de efectuados os quatro passos de Kimball, tem-se o seguinte modelo em estrela:

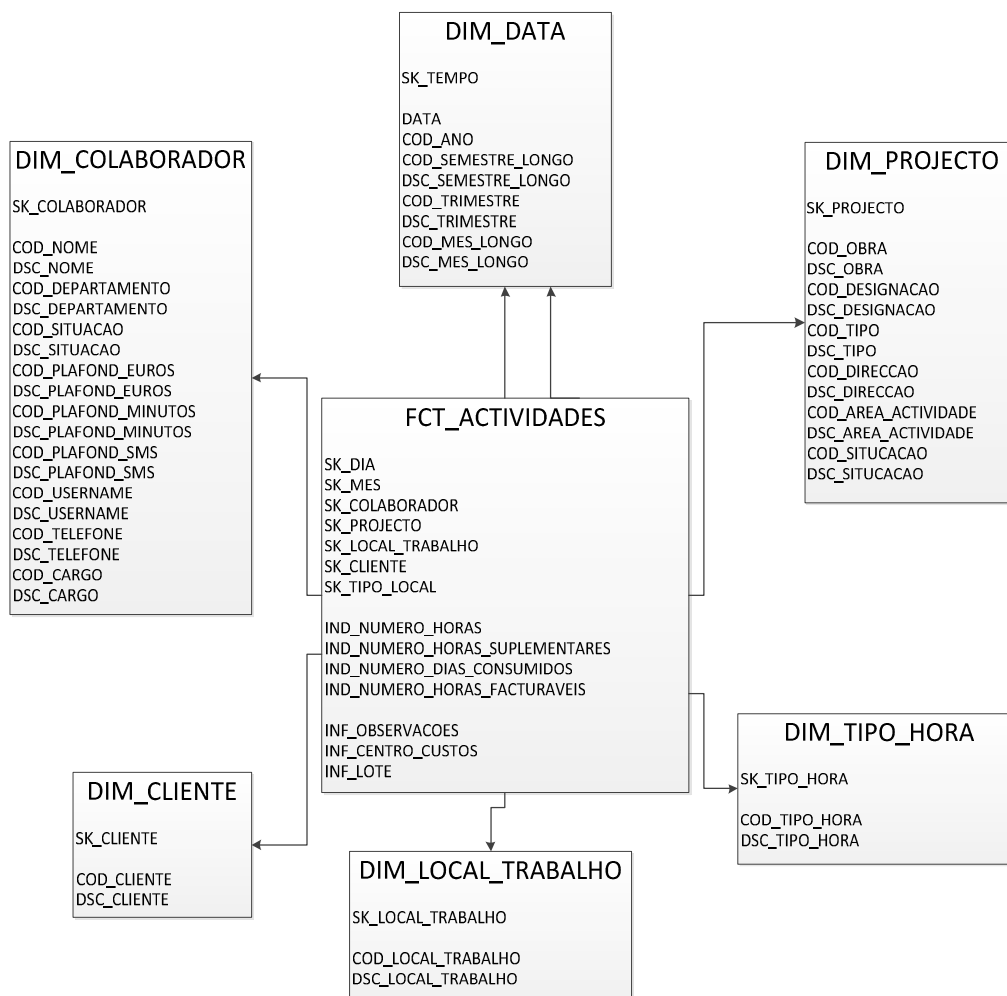


Figura 27 - Diagrama dimensional do sub-modelo Actividades - modelo de gestão de projectos

Área de gestão de projectos – Actividades Facturáveis

- **Passo 1 – Escolher a área funcional**

Gestão de projectos – Actividades Facturáveis

- **Passo 2 – Declarar o grão**

O número de horas facturáveis de um dado colaborador num determinado dia.

- **Passo 3 – Identificar as dimensões**

Dimensão colaborador, e tempo.

- **Passo 4 – Identificar os factos**

Número de horas facturáveis.

Para esta sub-área de gestão de projectos e depois de efectuados os quatro passos de Kimball, tem-se o seguinte modelo em estrela:

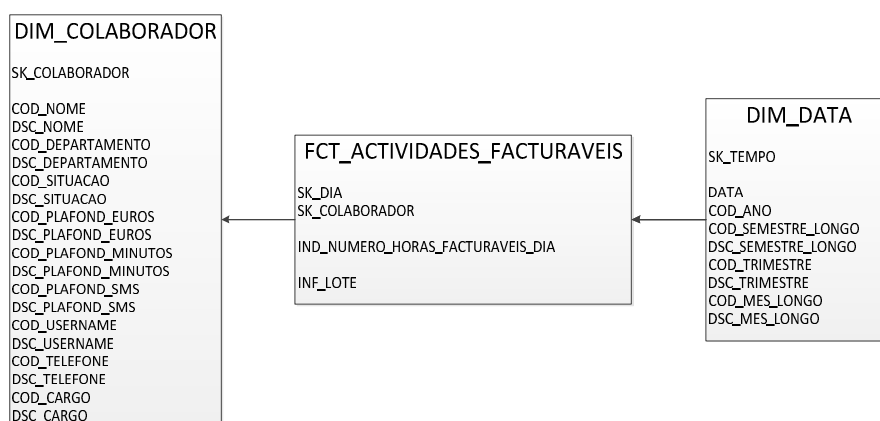


Figura 28 - Diagrama dimensional do sub-modelo Actividades Facturáveis - modelo de gestão de projectos

Área de gestão de projectos – Utilização de mão-de-obra

- **Passo 1 – Escolher a área funcional**

Gestão de projectos – Utilização de mão-de-obra

- **Passo 2 – Declarar o grão**

Uma linha da tabela de factos representa a capacidade máxima instalada, a capacidade disponível, a capacidade utilizada/ocupada e a capacidade facturada por colaborador num determinado mês.

- **Passo 3 – Identificar as dimensões**

Dimensão colaborador, e tempo.

- **Passo 4 – Identificar os factos**

Capacidade máxima instalada, capacidade disponível, capacidade utilizada/ocupada e capacidade facturada.

Para esta sub-área de gestão de projectos e depois de efectuados os quatro passos de Kimball, tem-se o seguinte modelo em estrela:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

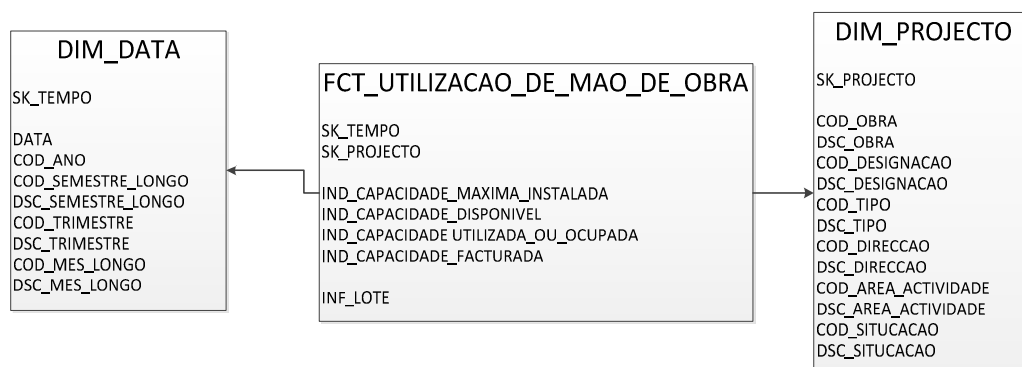


Figura 29 - Diagrama dimensional do sub-modelo Utilização de mão-de-obra - modelo de gestão de projectos

Área de gestão de projectos – Targets

- **Passo 1 – Escolher a área funcional**

Gestão de projectos – *Targets*

- **Passo 2 – Declarar o grão**

Uma linha da tabela de factos representa o *target* a atingir por colaborador.

- **Passo 3 – Identificar as dimensões**

Dimensão colaborador

- **Passo 4 – Identificar os factos**

Target.

Para esta sub-área de gestão de projectos e depois de efectuados os quatro passos de Kimball, tem-se o seguinte modelo em estrela:

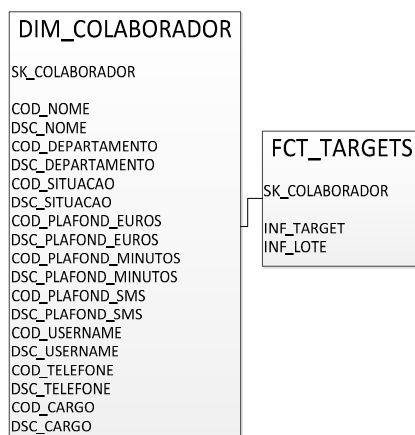


Figura 30 - Diagrama dimensional do sub-modelo Targets - modelo de gestão de projectos

4.4.5 Matriz de validação de métricas

A matriz de validação de métricas para o modelo de gestão de projectos e os sub-modelos a si associados, encontra-se na seguinte tabela:

| Gestão de projectos | Dimensões | | | | | | | | | | |
|---------------------------------|-----------|------|-------------|---------|-----------------|---------------------|----------|---------|-------------------|--------------|----------|
| | Tipo | Data | Colaborador | Periodo | Tipo de chamada | Tipo de comunicação | Projecto | Cliente | Local de trabalho | Tipo de hora | Incluir? |
| Métricas elementares | | | | | | | | | | | |
| Número de dias consumidos | # | x | x | | | | x | x | x | x | Sim |
| Número de horas facturáveis | # | x | x | | | | x | x | x | x | Sim |
| Número de horas | # | x | x | | | | x | x | x | x | Sim |
| Número de horas suplementares | # | x | x | | | | x | x | x | x | Sim |
| Número de horas facturáveis dia | # | | | | | | x | x | x | x | Sim |
| Target | # | | x | | | | | | | | Sim |
| Métricas agregadas | | | | | | | | | | | |
| Capacidade máxima instalada | ∑ | x | x | | | | | | | | Sim |
| Capacidade disponível | ∑ | x | x | | | | | | | | Sim |
| Capacidade utilizada/ocupada | ∑ | x | x | | | | | | | | Sim |
| Capacidade facturada | ∑ | x | x | | | | | | | | Sim |
| Métricas derivadas | | | | | | | | | | | |
| Taxa de disponibilidade | % | x | x | | | | | | | | Não |
| Taxa de ocupação | % | x | x | | | | | | | | Não |
| Taxa de facturação | % | x | x | | | | | | | | Não |

Tabela 7 - Matriz de validação de métricas do modelo de gestão de projectos

4.5 Physical Design

“This chapter outlines the steps required to turn the logical design into a physical database” (Kimball *et al.* 2008).

4.5.1 Bases de dados

Para este caso de estudo, e seguindo esta arquitectura e nomenclatura descrita em 4.3, tem-se as bases de dados: **BI4ALL_SA**, **BI4ALL_CONFIG**, **BI4ALL_DW** e **BI4ALL_LOG**.

Para visualização gráfica dessas bases de dados, consultar o **Anexo D – Bases de dados**.

Foram criadas para este trabalho tabelas como prefixo **MAP** (na base de dados BI4ALL_CONFIG) e representam a forma encontrada para combater um problema de dados que vinha do sistema fonte Optimus, mais propriamente o ficheiro excel fornecido por essa organização. A coluna período aí contida, continha várias definições diferentes para o mesmo conceito. Um exemplo real era por exemplo ter-se para a descrição do período “Fim-de-semana”, no Excel vinha umas vezes “fim de semana”; “fim semana”, “fim de seman”, etc. Isto obrigou a criar ao nível do pacote de ET, um processo de *fuzzy logic* que permitisse ao longo do tempo “adivinhar” com base no que já tinha armazenado, qual a verdadeira tradução para o período contido no novo ficheiro. A **MAP_NOVOS_PERIODOS** armazena todas as novas descrições de períodos que ainda não existiam na **MAP_SINONIMOS_PERIODOS**, dando uma sugestão para essa nova descrição e armazenado-a na tabela **MAP_NOVOS_PERIODOS**.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Na BI4ALL_DW estão contidas as *views* que fazem a ponte entre o DW e a solução OLAP SSIS. Numa perspectiva mais técnica, é nas *views* que a linha *dummy* é inserida nas dimensões. Deste modo, sempre que não é possível uma linha da factual cruzar com a dimensão, essa linha da factual ficará a apontar para a chave substituta inválida – Não Disponível (ND) que é criada apenas na *view* e não nas tabelas **DIM_Dimensao**. A nível de ETL, ao ocorrer esse cruzamento entre a factual e as dimensões e, dando-se esse caso de falha (ou seja, não existe correspondência entre o registo da factual com um registo da dimensão) a chave estrangeira na factual irá tomar o valor *null*. Ao fazer isso, e na execução das *views*, no caso das factuais, se a chave estrangeira for *null*, ela irá ser instanciada com o valor “-1”. Na execução da *view* das dimensões, a linha *dummy* irá ter a chave substituta com o valor “-1”, sendo então para esta que o registo na factual irá apontar na dimensão. Todas as chaves substitutas são criadas em tempo real aquando a inserção de registos nas dimensões. Isso é feito automaticamente pelo gestor de base de dados do SQL Server, em que no campo respectivo à chave substituta, se define a propriedade *Auto Identity* igual a verdadeiro, sendo incrementado de um em um sempre que um novo registo é inserido. Para um exemplo do código de criação uma *view*, verificar o **Anexo E - Exemplo do código de criação de uma *view***.

4.6 ETL Design & Development

Depois de se ter definido a camada física, segue-se a fase de ETL dos dados no DW. Para isso, utilizou-se a tecnologia SSIS no desenvolvimento de cada pacote. Seguiu-se a arquitectura ET+L em termos de desenvolvimento e apresentação dos pacotes.

4.6.1 Configurações nos pacotes

Em cada pacote, seja de ET, ou de L, têm-se configurações ao nível do pacote de variáveis que contêm os seus valores armazenados na tabela de configurações na base de dados de *Config*, apresentada anteriormente. Um exemplo dessas configurações nos pacotes de SSIS encontra-se ilustrado na seguinte figura:

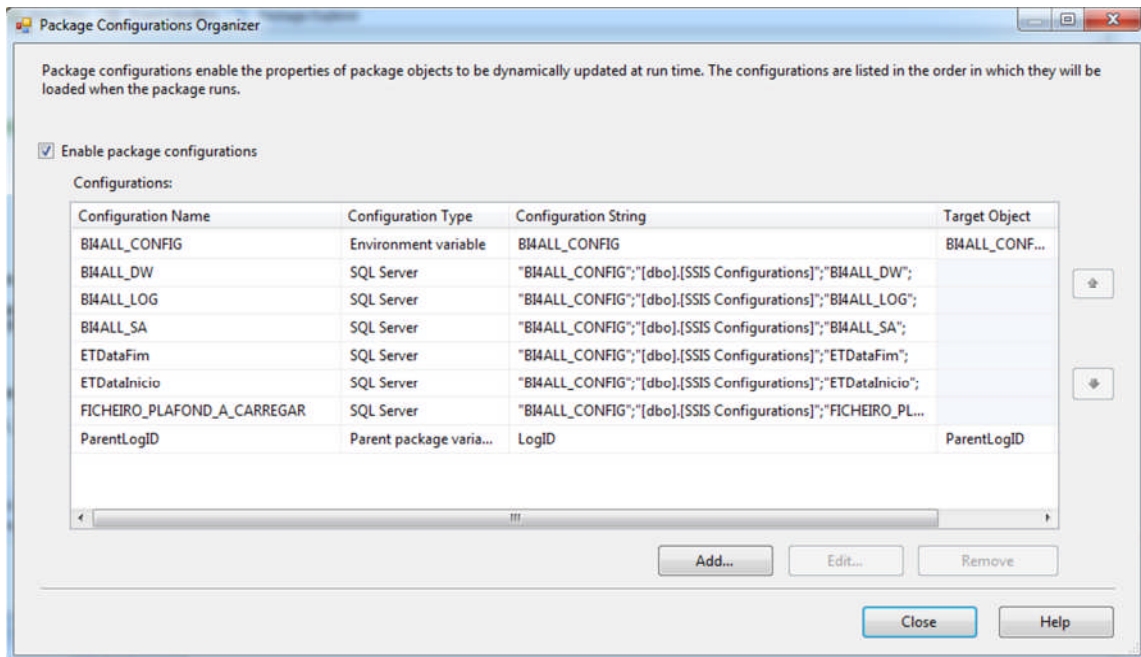


Figura 31 - Configuração de pacotes no SSIS

Como se consegue observar pela imagem acima, é neste organizador de configurações ao nível do pacote, que todos os caminhos para as bases de dados, ficheiros, etc, são definidos. A principal configuração a expôr aqui inicialmente é a “BI4ALL_CONFIG” – esta é a responsável por indicar o caminho onde se encontra a base de dados de *config* no servidor onde a solução está alojada; isto é efectuado graças a uma variável de ambiente que é criada na máquina em questão, apontando para a base de dados de *config* (neste caso, a BI4ALL_CONFIG exposta no ponto 4.5). A título de exemplo, tal variável ambiente poderia conter este valor como apontador: Provider=SQLNCLI10.1;DataSource=.;IntegratedSecurity=SSPI;InitialCatalog=BI4ALL_CONFIG.

As outras configurações, alojadas na base de dados de *config*, contêm os caminhos para as bases de dados em questão – como se pode ver na figura acima, estão aí armazenados os caminhos para a base de dados BI4ALL_CONFIG, BI4ALL_DW, etc. Também se consegue indicar qual o caminho onde estão alojados ficheiros a carregar no pacote de SSIS em questão.

A vantagem desta abordagem centra-se na portabilidade que se consegue dar ao projecto. Em vez de definir em cada pacote, de uma forma individual, quais são os caminhos para cada uma das fontes de dados ou directórios de destino, isso é apenas definido uma vez e armazenado num repositório único, onde todos os pacotes acedem para obter os caminhos necessários – apenas se tem que mudar a variável de ambiente se mudarmos a solução fisicamente de lugar; ao passar a estar num novo servidor, altera-se o caminho para a base de dados de *config* da solução, e automaticamente cada pacote vai conseguir ligar-se com sucesso à tabela de configurações e

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

aceder a tudo a que necessita para ocorrer o processamento com sucesso. O esforço na mudança é praticamente nulo.

4.6.2 Extração e transformação

Para a realização da fase de ET dos dados contidos nos sistemas fonte, tem-se a seguinte figura que ilustra qual a estrutura e nomenclatura adoptada para todos os pacotes de desenvolvimento desta fase, tendo-se utilizado o SSIS como referido em 4.3 e ilustrado na seguinte figura:

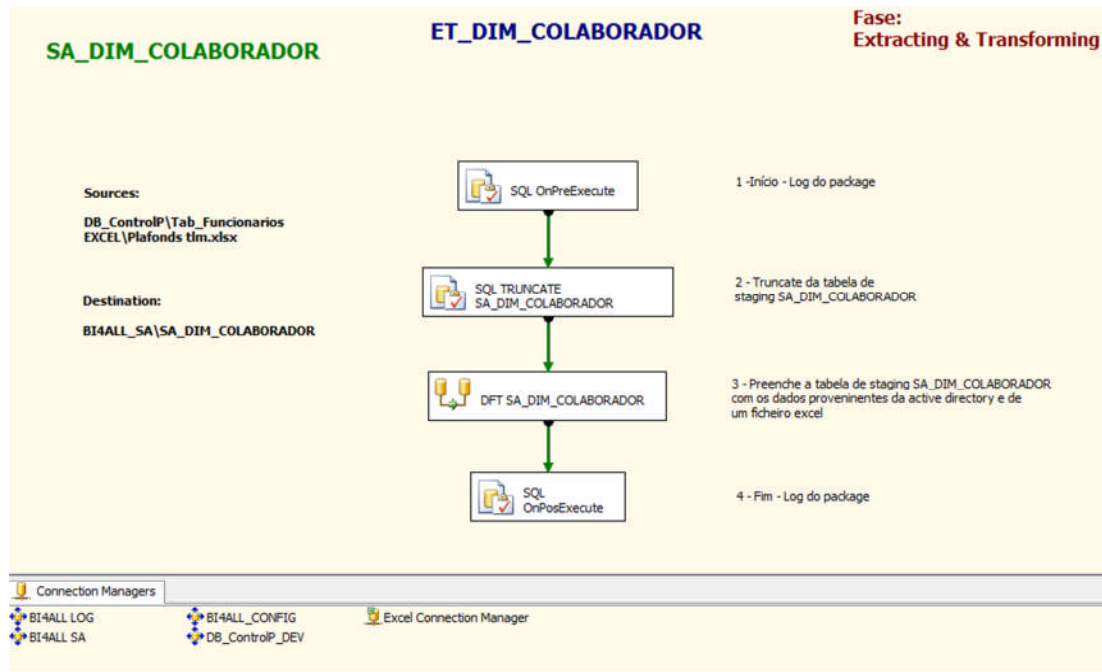


Figura 32 - Pacote de ET relativo à dimensão Colaborador

A figura anterior mostra o que está contido no *Control Flow* deste pacote. Temos no cabeçalho a identificação de qual o pacote que se está a tratar e a fase relativa ao ETL (se à fase de ET ou se à fase de L), tendo-se assim uma arquitectura do tipo ET+L. Isto permite não só uma melhor forma de se organizar todo o desenvolvimento, como isolar a fase de extração e limpeza dos dados, da fase de carregamento para as dimensões e factuais e facilitar acções de evolução do sistema.

Ainda sobre a figura acima, tem-se do lado esquerdo quais os sistemas fonte aos quais se vai retirar os dados necessários para a dimensão/factual em causa e também qual o destino que os dados irão ter depois de transformados. Na área central tem-se os componentes que vão tratar de todo o processo e, na parte mais inferior da figura, tem-se as conexões necessárias para o pacote ser processado com sucesso.

O primeiro componente com a descrição *SQL OnPreExecute* irá armazenar na base de dados de *log*, mais propriamente na tabela *BI4ALL_LogExecucao*, informação relativa sobre qual o

pacote que foi executado, por quem foi executado, a que horas, etc. O segundo componente, *SQL Truncate SA_DIM_COLABORADOR*, irá realizar um *delete* à tabela de *staging* em causa. As tabelas de *staging* neste projecto são sempre truncadas, sejam tabelas *staging* de dimensões como de factuais (estas tabelas estão armazenadas na já referida *staging area*). O terceiro componente é relativo ao processo de ET em si – é denominado como o *data flow task*. Este componente, relativo ao pacote ET_DIM_COLABORADOR, está ilustrado na seguinte figura:

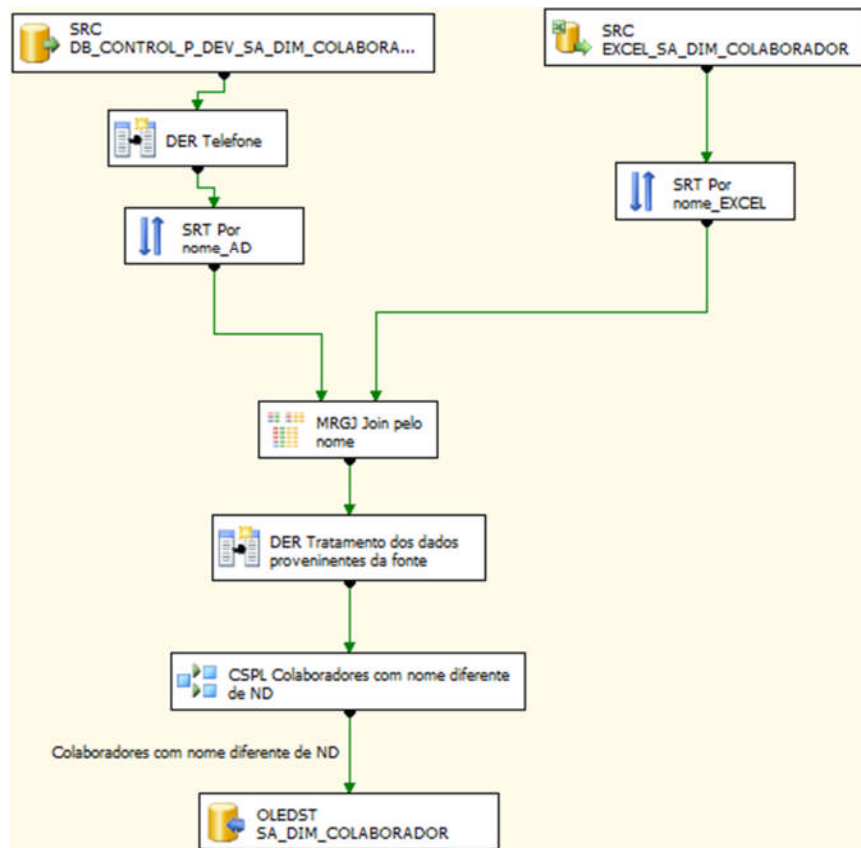


Figura 33 - Fluxo de dados do pacote de ET da dimensão Colaborador

Consegue-se então observar que se começa por extrair todos os dados necessários dos sistemas fonte, sendo neste caso uma das fontes uma base de dados em *SQL Server* e a outra uma folha de cálculo excel. Depois dessa extração, todos os dados são transformados (neste caso nos componentes com o prefixo DER), cruzados (neste caso através do componente com o prefixo MRGJ) e depois entregues no destino, isto é, a base de dados de *staging* (neste caso, a tabela temporária SA_DIM_COLABORADOR).

Relativamente ainda à figura 32, tem-se o último componente, *SQLOnPosExecute*, que irá armazenar informação na tabela de *log* relativo ao sucesso de execução do pacote, como o tempo final de execução.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Para a fase de ET das factuais, a lógica seguida no ET das dimensões mantém-se, sendo a estrutura semelhante em termos de nomenclatura e de componentes usados. Para este caso de exemplo, tem-se a apresentação da fase ET do pacote relativo à factual das chamadas efectuadas pelos colaboradores e cujo *control flow* se encontra na seguinte figura:

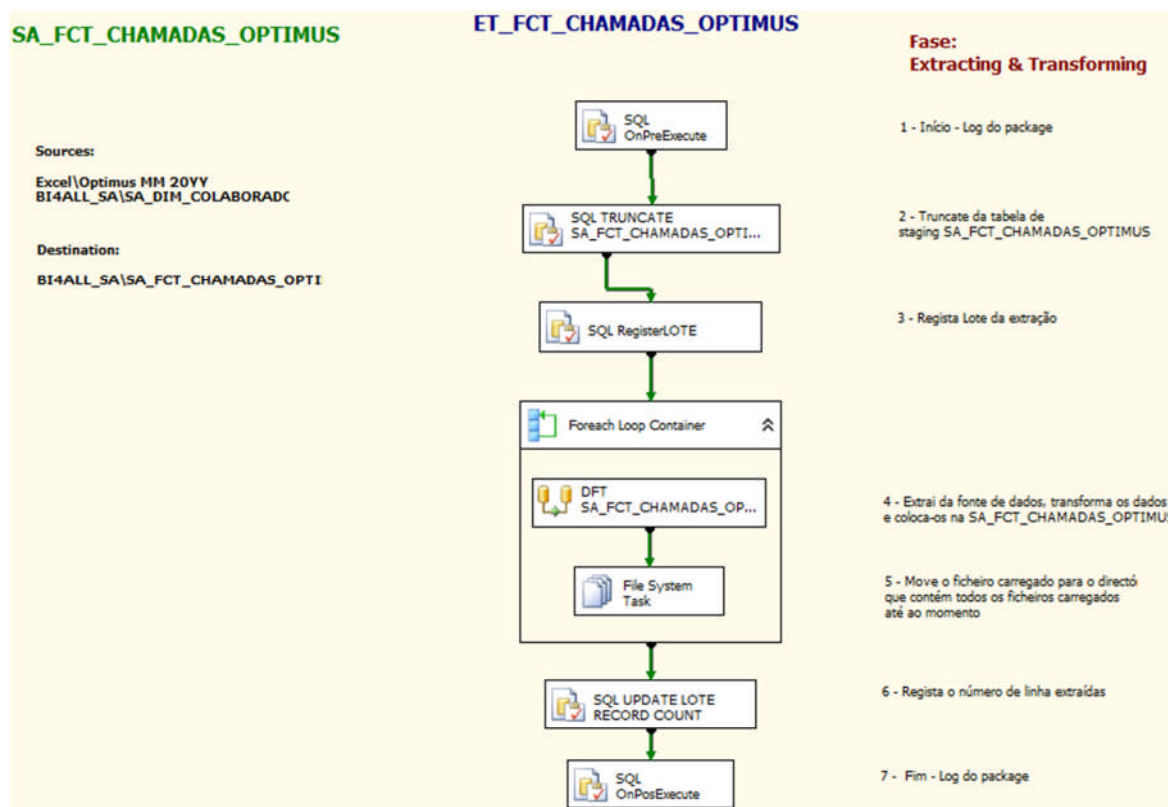


Figura 34 - Pacote de ET relativo à factual das chamadas do modelo de controlo e redução de custos

Como se consegue constatar pela figura acima, a estrutura é semelhante ao *control flow* da dimensão colaborador, diferindo apenas em três componentes: *Foreach Loop Container* – este componente serve para enquanto houver ficheiros num certo directório, o componente irá lê-los todos; *File System Task* – este componente vai no fim de processar cada ficheiro nesse directório estabelecido anteriormente, movê-lo para um directório de “ficheiros carregados”; e, o *SQL UPDATE LOTE RECORD COUNT* – este componente irá preencher na factual na coluna lote, o lote relativo a esta extração (a cada novo carregamento, esta coluna é incrementada em uma unidade) e registar o número de linhas extraídas. Os outros componentes têm o mesmo papel que o explicado acima no ET_DIM_COLABORADOR. Sobre o componente de *DFT SA_FCT_CHAMADAS_OPTIMUS*, é particular para este caso e que se encontra ilustrado na seguinte figura:

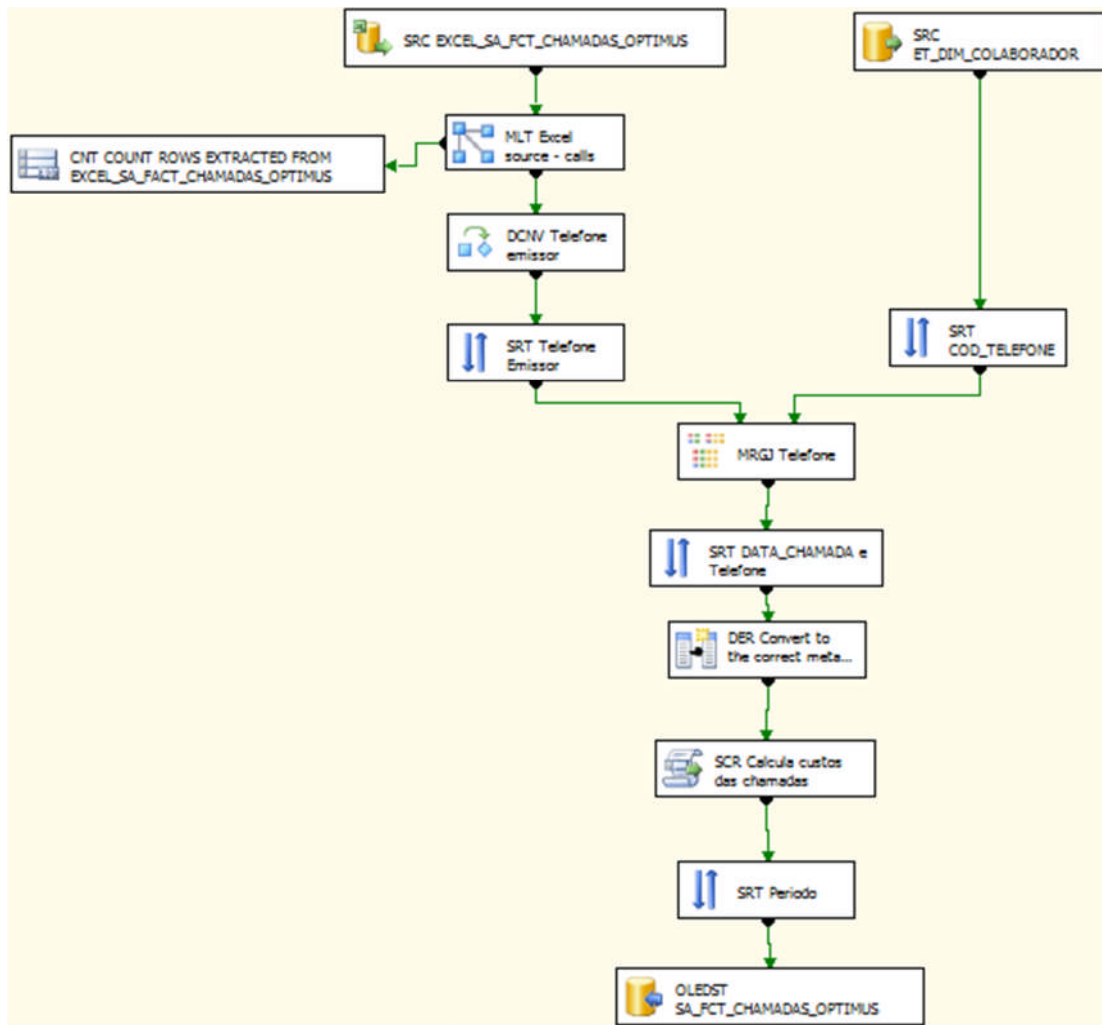


Figura 35 - Fluxo de dados (data flow) do pacote de ET da factual das chamadas do modelo de controlo e redução de custos

Similarmente ao *data flow task* seguido no ET_DIM_COLABORADOR, todos os dados são extraídos ao início. De realçar neste ponto o componente *SCR Calcula custos das chamadas*, que é um *script* em C# e que está responsável pela atribuição de custos das chamadas realizadas de acordo com o *plafond* definido para cada um dos colaboradores da BI4ALL.

Consegue-se então observar que se começa por extrair todos os dados necessários dos sistemas fonte, sendo neste caso uma das fontes uma base de dados em *SQL Server* e a outra uma folha de cálculo excel. Depois dessa extração, todos os dados são transformados (neste caso nos componentes com o prefixo DER), cruzados (neste caso através do componente com o prefixo MRGJ) e depois entregues no destino, isto é, a base de dados de *staging* (neste caso, SA_DIM_COLABORADOR).

4.6.3 Carregamento

Para a realização da fase de carregamento dos dados contidos na área de *staging*, também já referido neste documento como a fase de **L**, tem-se a seguinte figura que ilustra qual a estrutura e nomenclatura adoptada para todos os pacotes de desenvolvimento desta fase (pacotes relativos ao carregamento de dimensões):

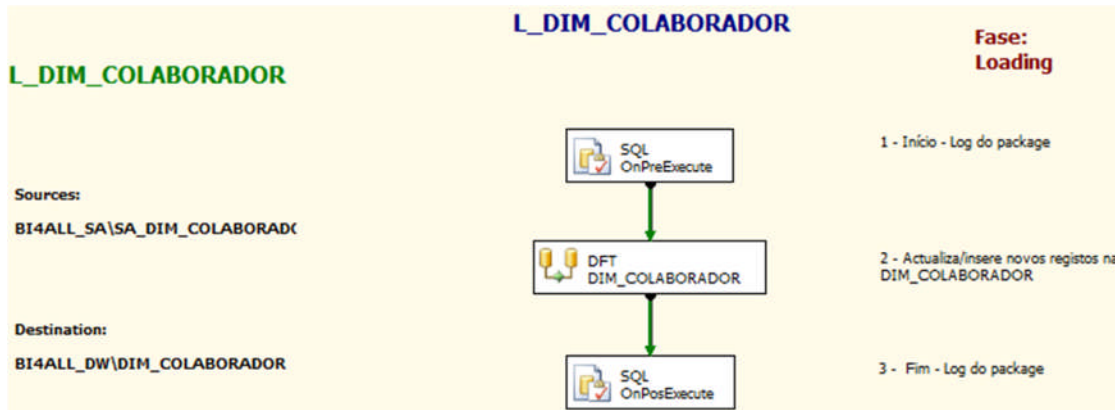


Figura 36 - Pacote de L relativo à dimensão Colaborador

Como se consegue observar, é extremamente similar com o pacote relativo à fase de ET, mas os componentes relativos ao truncamento das tabelas de *staging* já não se encontram aqui, sendo todo o processo de carregamento efectuado no componente *data flow task* DFT DIM_COLABORADOR, e ilustrado na seguinte figura:

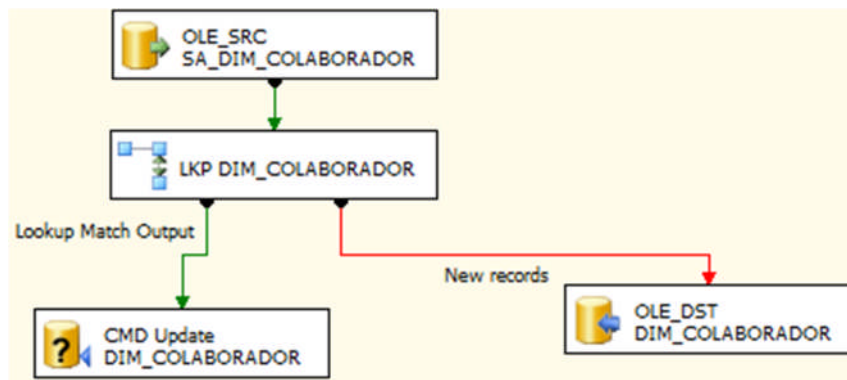


Figura 37 - Fluxo de dados (data flow) do pacote de L da dimensão Colaborador

O que é feito no carregamento de uma dimensão, neste caso da dimensão colaborador, é ter-se uma fonte de dados que aponta para a tabela de *staging* em causa e que já contém todos os dados tratados na fase de ET. Esses dados vão cruzar com o que já existe no DW através do componente *Lookup* (LKP DIM_COLABORADOR) - se existirem dados que venham da *staging area* e que cruzem com a dimensão no DW, é realizada uma actualização ao DW e é

seguido o fluxo *Lookup Match Output*. Em caso de serem novos registos a serem inseridos na dimensão, esses seguem o fluxo *New records* e são então inseridos na tabela de destino.

Para a fase de L das factuais, a lógica seguida no L das dimensões altera-se significativamente, pois em vez de ser realizado um *lookup* com a factual no DW, vão ser efectuados vários *lookups* mas com cada dimensão que faz parte do grão da factual em causa. A estrutura é também ligeiramente diferente comparativamente com o L das dimensões, tendo-se então como exemplo o pacote de carregamento da factual das chamadas – L_FCT_CHAMADAS_OPTIMUS:

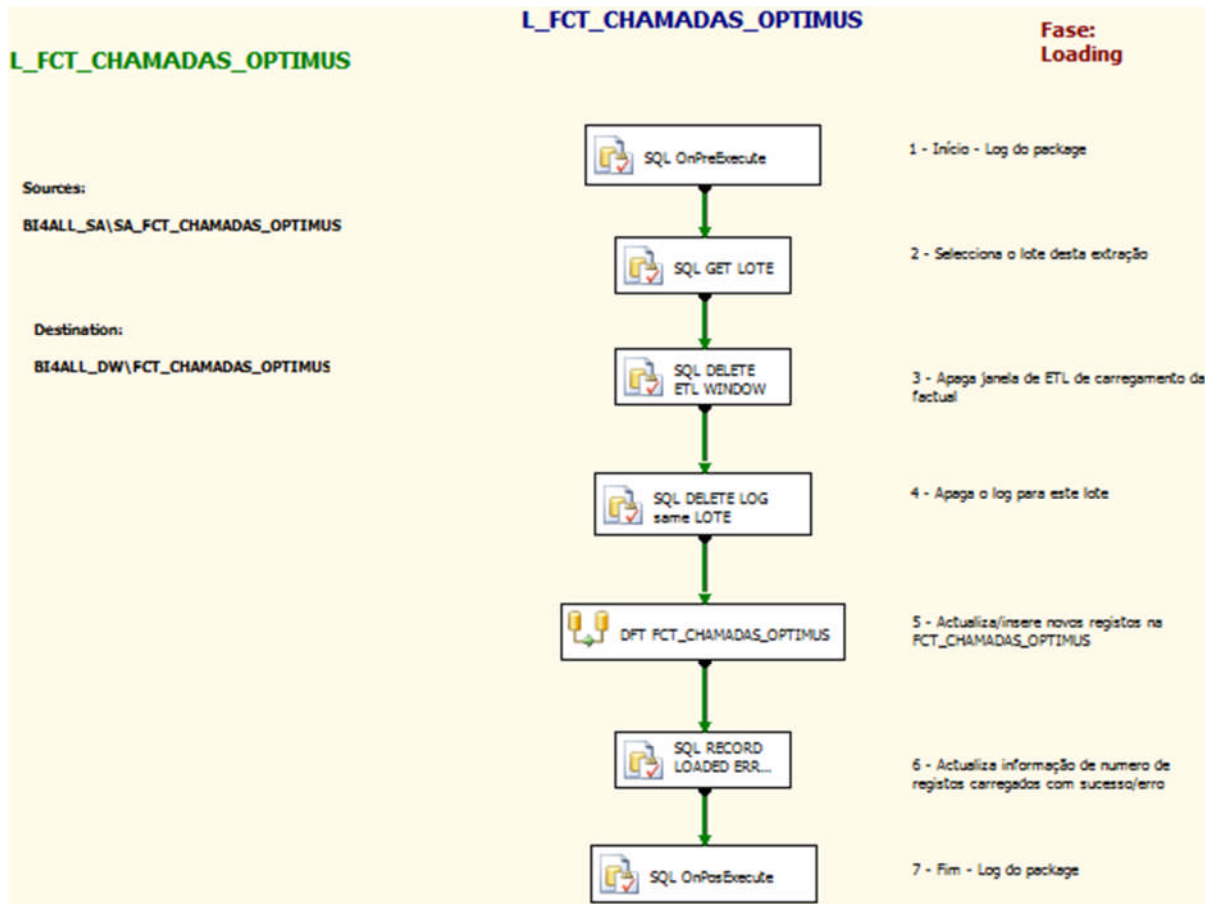


Figura 38 – Pacote de L relativo à factual das chamadas do modelo de controlo e redução de custos

Observando a figura acima, no carregamento das factuais existem mais alguns componentes quando comparado com o pacote de carregamento das dimensões. Começa por se diferenciar com o componente de *SQL Task*, nomeadamente o *SQL GET LOTE*, que vai buscar qual o lote relativo a esta extração, e de seguida o *SQL DELETE ETL WINDOW* que irá apagar na factual contida no DW a janela de carregamento a ser efectuada. Por exemplo, se estiver a ser carregado dados relativos ao mês de janeiro de 2011 na tabela de *staging* da factual, o mesmo período vai ser apagado na factual do DW para que os novos dados sejam inseridos e não ocorra duplicação

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

de registos para o mês em causa; caso seja um novo mês a ser carregado, o comando de *delete* não irá ter qualquer efeito. Isto é muito melhor em termos de desempenho do que fazer actualizações a milhões e milhões de linhas (o que pode ser o caso, quando o período temporal é muito grande por exemplo) quando se está a carregar novamente dados (histórico), que por alguma razão em específico, tivesse que ser reprocessado novamente no DW. No seguinte componente, *SQL DELETE LOG same LOTE*, vai-se proceder ao apagar dos registos que estão contidos na tabela de *log*, neste caso da tabela LOG_FCT_CHAMADAS_OPTIMUS, com o último valor de lote contido na tabela de *staging*. Isto serve para que ao reprocessar a fase de carregamento desta factual, que se apague o que foi armazenado na tabela de log relativo à extração anterior, de forma a apenas registar aquilo que falhou em termos de processamento apenas relativamente ao último lote de extração da fase de ET para esta factual. Depois, chega a fase de carregamento da factual no DW através do componente de *data flow task* DFT FCT_CHAMADAS_OPTIMUS, que está representado na seguinte figura:

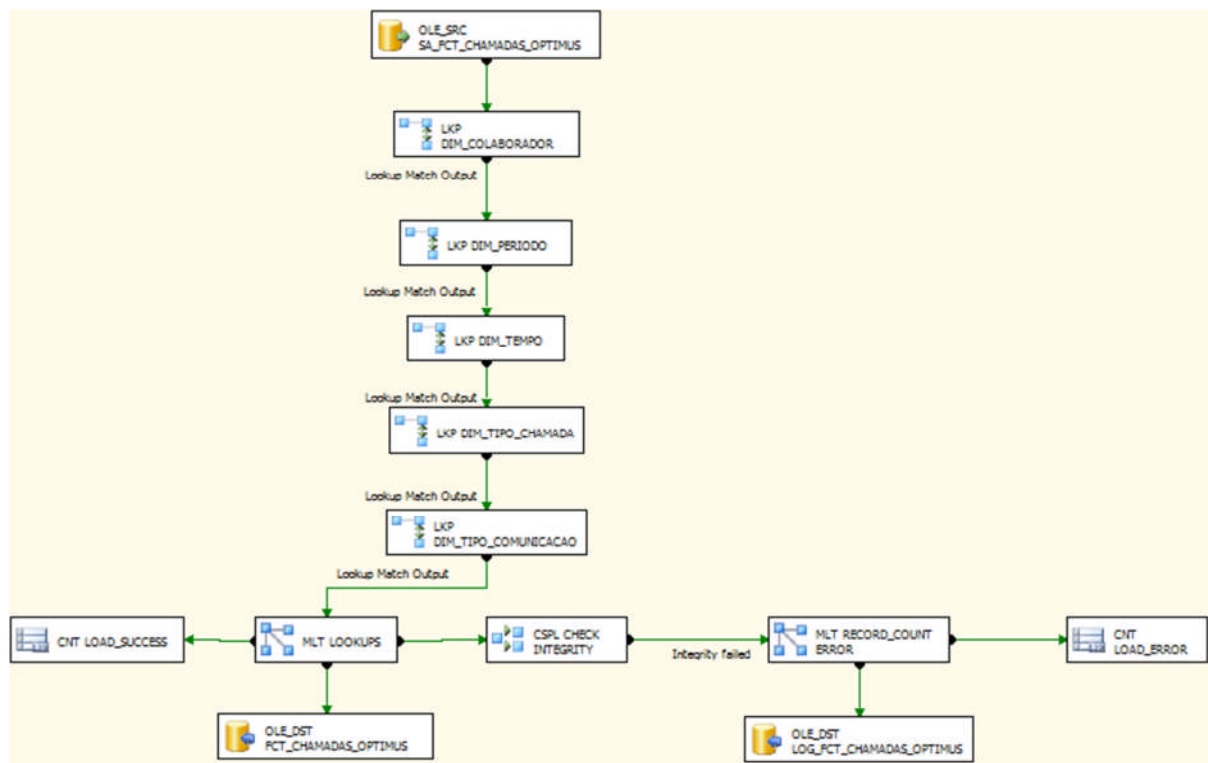


Figura 39 - Fluxo de dados (*data flow*) do pacote de L da factual das chamadas do modelo de controlo e redução de custos

Data Flow L_FCT_CHAMADAS

Visualizando a figura 29, tem-se que como ponto de partida ao carregar a factual contida no DW, um acesso à tabela de *staging* da factual em causa (SA_FCT_CHAMADAS_OPTIMUS). Depois, cada linha aí contida, irá cruzar com todas as dimensões através de *lookups* – com isto

irá obter-se as chaves substitutas que estão nas dimensões como chaves estrangeiras na factual. Em caso de não haver correspondência entre a factual e uma dimensão, o pacote não falha, deixando assim passar o registo com a chave substituta igual a *null*. Para os casos que falham, são tratados nas *views* onde cada registo a *null* passa a obter o valor -1 (valor definido como “Não Aplicável” nas dimensões, já mencionado anteriormente neste documento).

Depois do último *lookup* presente na imagem acima, tem-se um componente do tipo *multicast* que vai multiplicar as linhas em vários sentidos: um sentido vai para o componente que conta o número de linhas carregadas no total – *CNT LOAD SUCESS*; redirecciona para o componente que insere as linhas na factual contida no DW – *OLE_DST FCT_CHAMADAS_OPTIMUS*; e outro fluxo que vai para um componente de decisão – *CSPL CHECK INTEGRITY*. Este último vai determinar de todas as linhas, quais as que vão ser inseridas na tabela de *log* (qualquer uma que contenha a chave substituta igual a *null*). Por fim, é realizada uma contagem de todas as linhas com erro inseridas na base de dados de *log*, cuja informação será inserida graças ao componente presente na figura 38: *SQL RECORD LOADED ERROR*.

4.6.4 Processamento das dimensões e factuais

Neste ponto, vai ser apresentado a estrutura de processamento das dimensões e factuais que também é realizada pelo SSIS. Tem-se vários pacotes com especificidades diferentes, tendo sido expostos no ponto anterior. Para exemplificar neste documento, vai-se expôr o pacote relativo ao processamento de toda a componente de ETL (1), o pacote de processamento das dimensões (2), o pacote de processamento das factuais (3) e, o pacote de processamento OLAP das dimensões e cubo Optimus (4).

Para o pacote de processamento de toda a componente de ETL (gestão de projectos e controlo e redução de custos) tem-se a seguinte figura que ilustra o processamento de todos os pacotes de ET+L:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

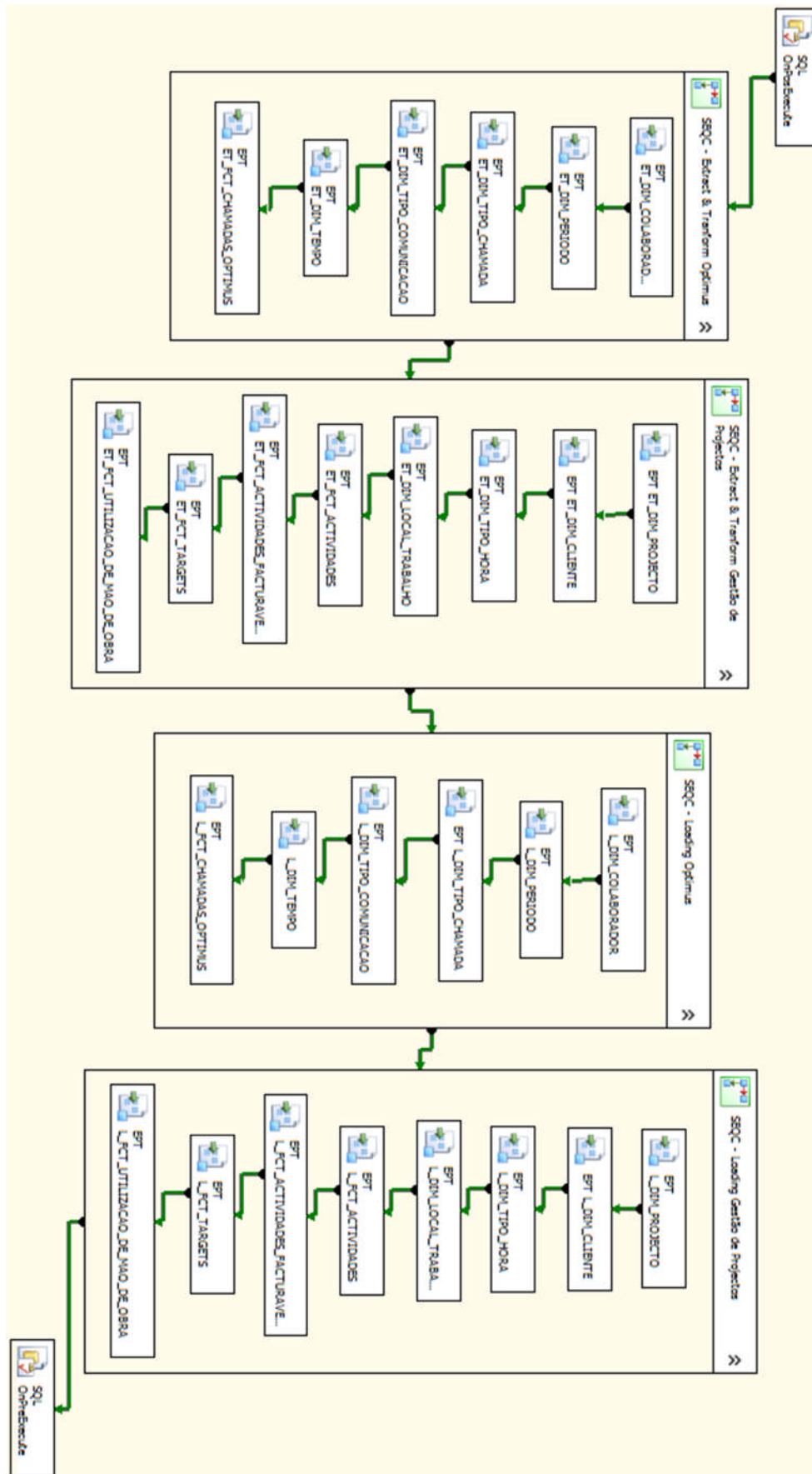


Figura 40 - Pacote relativo ao processamento global de todo o ETL

Como se pode constatar pela ilustração acima, todos os pacotes tanto de gestão de projectos como de controlo e redução de custos estão aí presentes, estando divididos logicamente por *sequence containers*, para uma melhor visualização e interpretação gráfica. Tanto dimensões como factuais estão aí presentes.

Tem-se agora um pacote apenas para procesamento de dimensões. A razão da existência de tal pacote recai na necessidade de por vezes apenas necessitarmos processar certos componentes do sistema (como por exemplo, processar apenas as dimensões), excusando-se assim de perder tempo de processamento em processar coisas desnecessárias para a necessidade existente no momento (senão apenas o pacote apresentado na figura 40 e o de processamento OLAP serviriam para resolver a questão). Tem-se então na seguinte ilustração o pacote de processamento das dimensões:

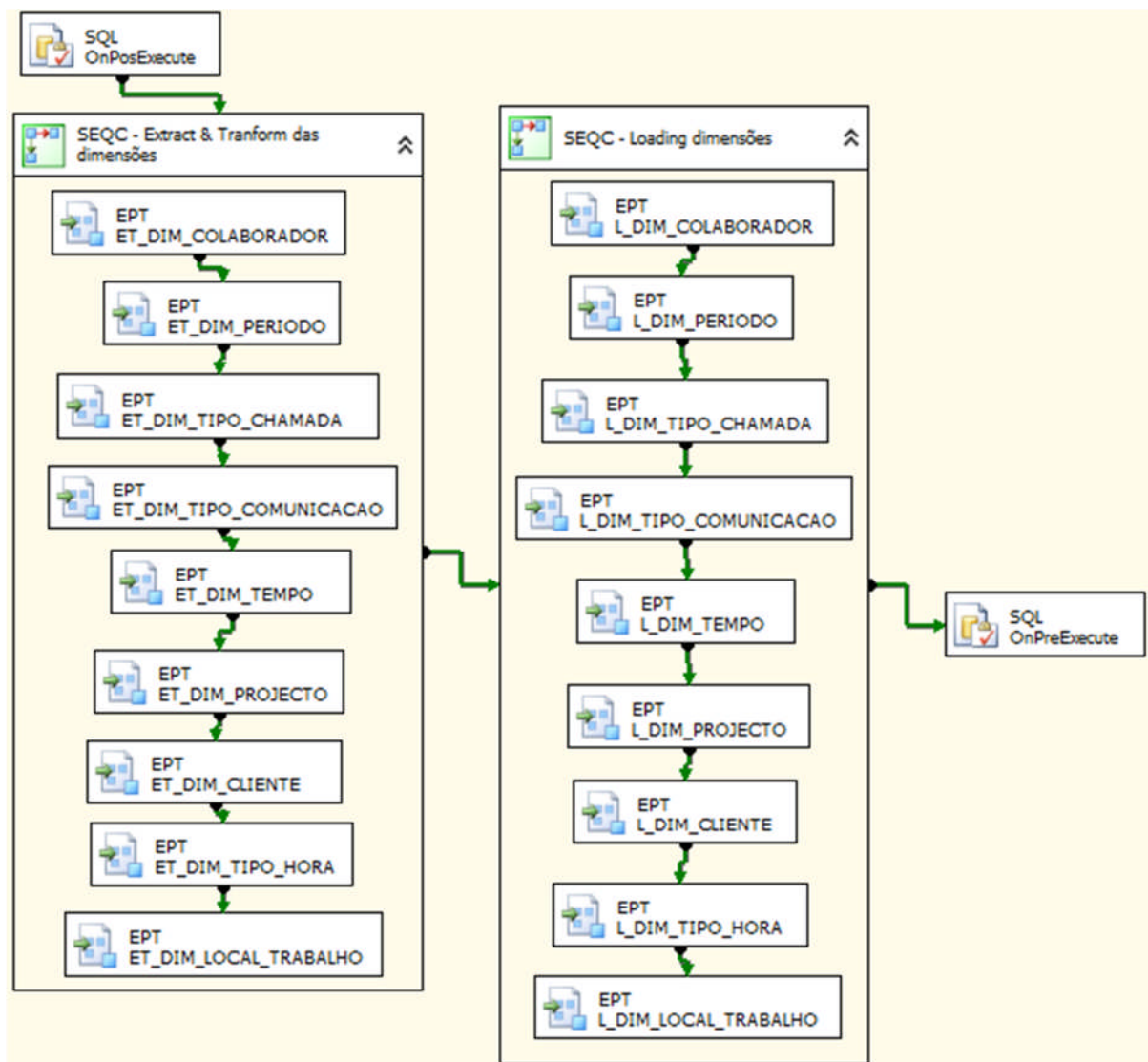


Figura 41 - Pacote de processamento das dimensões

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Como se pode observar, é bastante semelhante à ilustração do pacote de processamento global de todo o ETL, só aqui apenas se inseriu tudo em dois *sequence containers*, deixando as factuais de existir neste pacote.

Para o pacote de processamento das factuais a lógica é idêntica à anterior, e a razão da existência centra-se também da necessidade de por vezes ter que se processar componentes mais isoladamente e não o sistema global. Este pacote está ilustrado na seguinte figura:

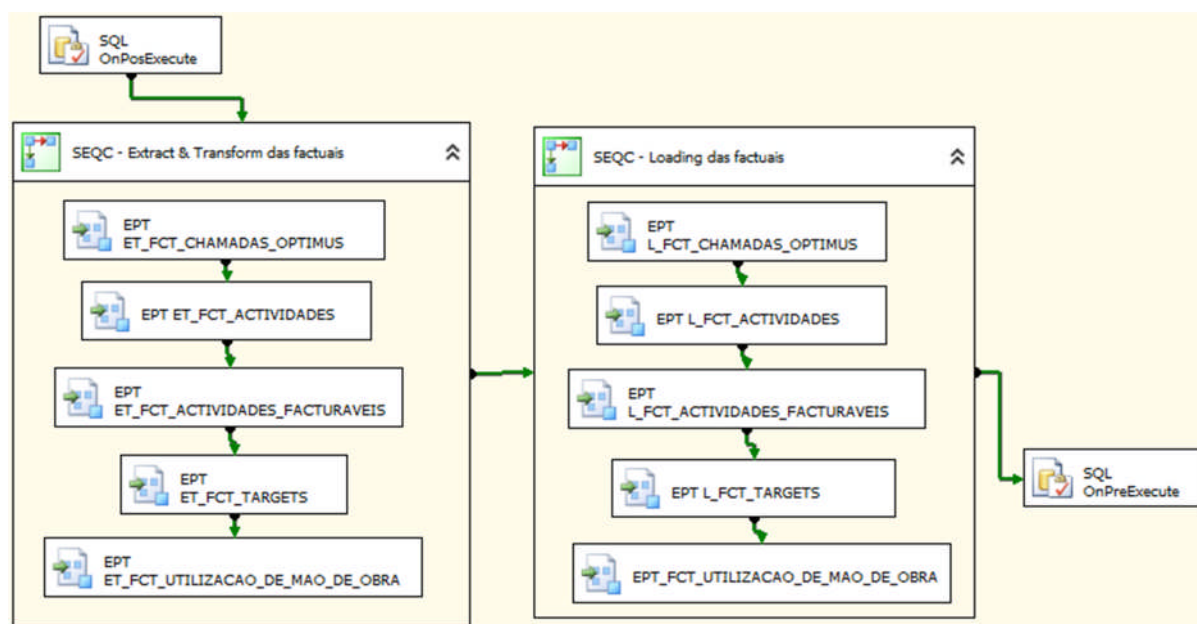


Figura 42 - Pacote de processamento das factuais

Como se consegue constatar, esta ilustração é bastante semelhante à anterior das dimensões, diferenciando-se apenas no aspecto deste processar apenas as factuais contidas no DW.

Por último, tem-se o pacote de processamento OLAP das dimensões e respectivo cubo – neste caso vai-se apresentar o cubo Optimus (controlo e redução de custos):

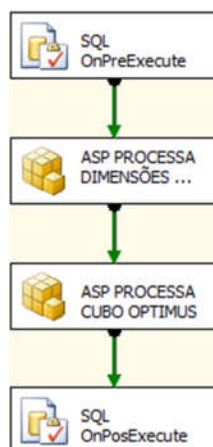


Figura 43 - Pacote de processamento OLAP

Na imagem acima identifica-se o fluxo necessário para processar as dimensões e respectivo cubo (neste caso, o cubo de controlo e redução de custos – também mencionado como cubo Optimus). Os componentes de *logging* (*SQL OnPreExecute* e *SQL OnPosExecute*) continuam presentes, sendo o processamento das dimensões realizado pelo componente *ASP PROCESSA DIMENSÕES* e o processamento do cubo pelo componente *ASP Processa CUBO OPTIMUS*. Existem dois modos de processamento que podiam ter sido adoptados tanto para o processamento das dimensões como para os cubos (1) “*Process Update*”: “Força uma re-leitura dos dados e realiza uma actualização aos atributos da dimensão (...) esta opção de processamento pode adicionar novos membros a uma dimensão e forçar uma re-leitura completa dos dados para uma actualização dos atributos desse objecto.” (Microsoft TechNet 2011). e (2) “*Process Full*”: “(...) Quando um *process full* é executado num objecto que já tenha sido processado, o *Analysis Services* apaga todos os dados nesse objecto, reprocessando-o de seguida (...)” (Microsoft TechNet 2011).

Para as dimensões, a escolha recaiu no “*Process Update*”, tal como ilustrado na seguinte figura:

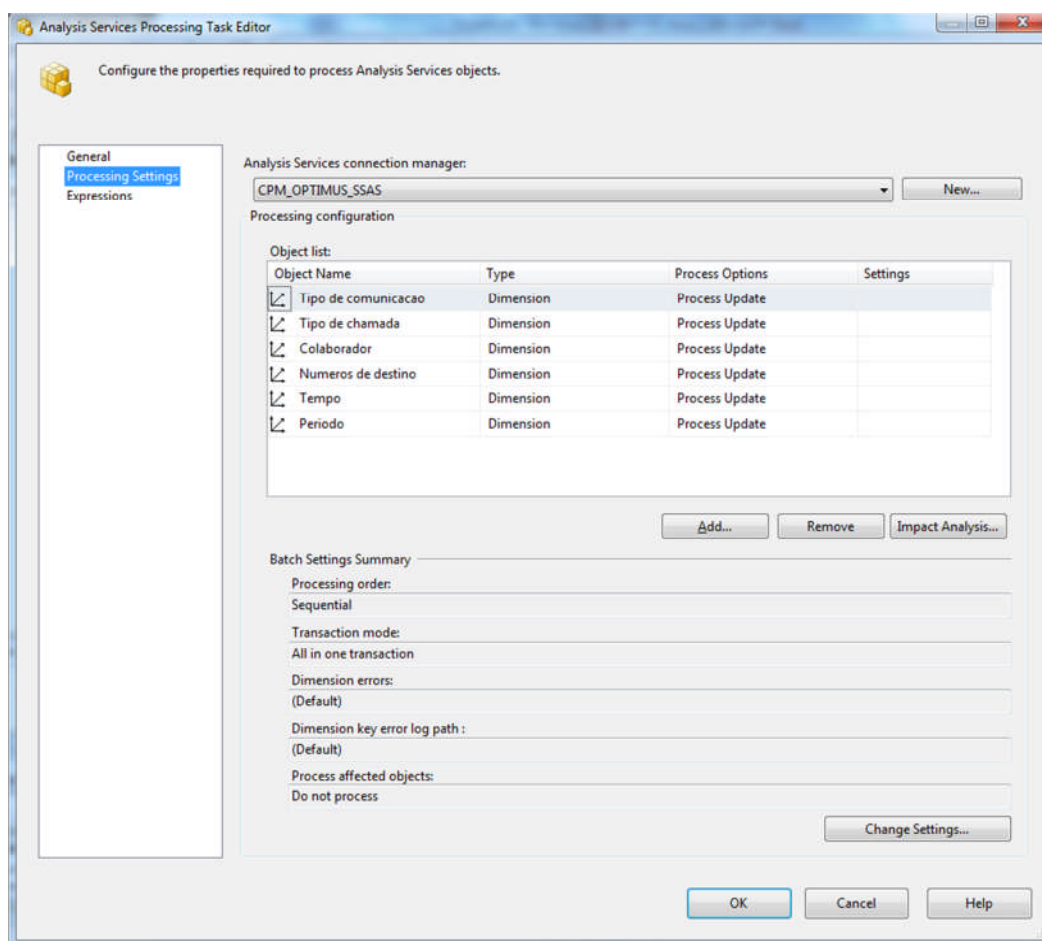


Figura 44 - Processamento de dimensões (*Analysis Services*)

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

A justificação centra-se em que numa dimensão não ser necessário estar a fazer *drop* aos dados aí contidos, pois as dimensões sofrem muito poucas alterações ao longo do tempo, sendo mais efectuado o processo de actualizações e, o número de re-leitura de linhas não é dispendioso pela razão mencionada anteriormente (poucas linhas nas dimensões quando comparado com as factuais). Já em termos de processamento de cubos, a opção de “*Process Full*” é a seleccionada pois re-ler todos os dados contidos numa factual (ou várias), pode ter um *performance* bastante prejudicial em termos de tempos de entrega de dados aos utilizadores. A opção de *drop* de todos os dados e processamento a seguir identifica-se como a melhor solução nestes casos, tendo sido adoptada para os dois cubos desenvolvidos para esta dissertação. A seguinte figura comprova o que foi mencionado anteriormente:

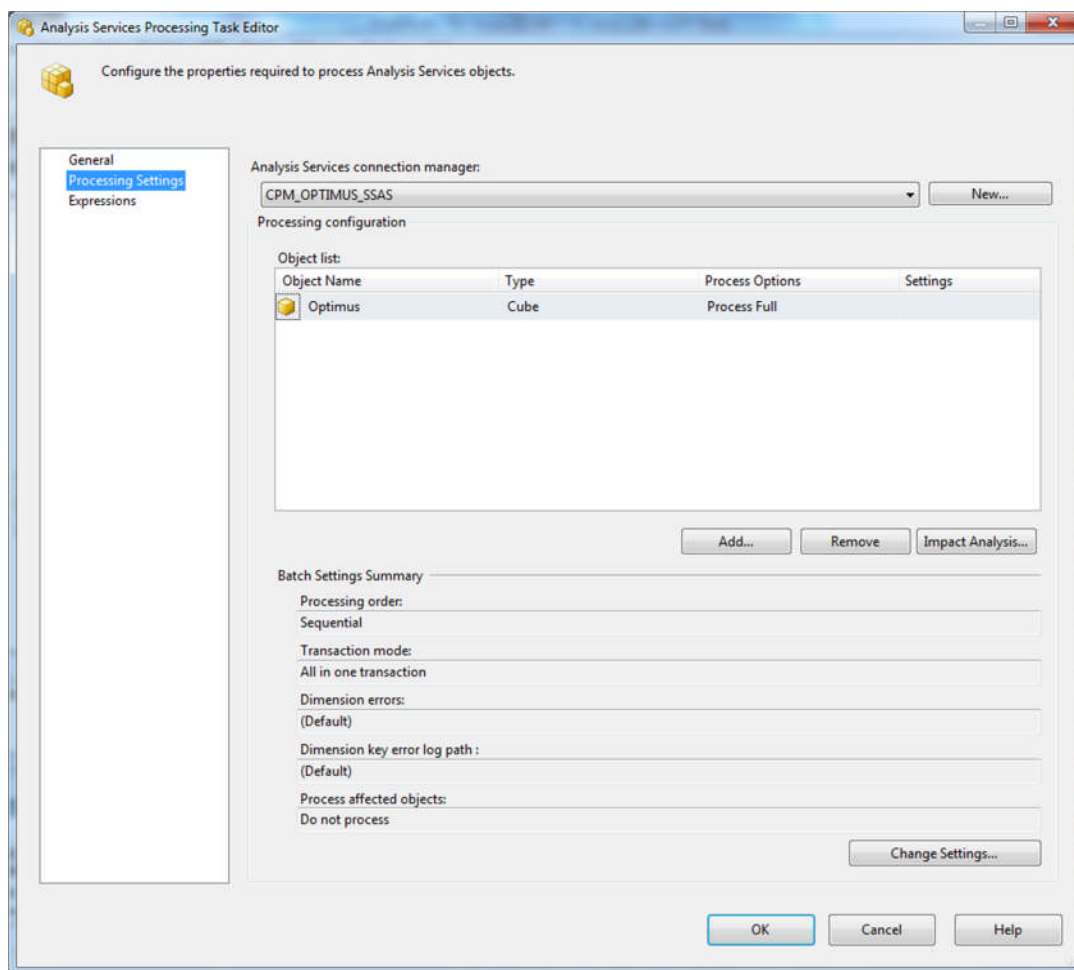


Figura 45 - Processamento do cubo do modelo de controlo e redução de custos

Para o processamento das dimensões e cubos se consiga realizar, é necessário criar uma solução em SSIS, onde toda a fase de exploração dos dados é realizada, sendo definidas dimensões e suas hierarquias, nomes de negócio, e os cubos com as suas métricas calculadas na fase de ETL e também as que são calculadas em tempo real ao se navegar no cubo (métricas como taxas são

um exemplo típicos de métricas calculadas apenas na fase de construção do cubo). Isto vai ser abordado **4.6.5**.

4.6.5 Solução Analysis Services – exploração dos dados

Para a exploração dos dados, como já foi referido ao longo do documento, foi utilizada a tecnologia SSIS. O que é feito neste ponto é a ligação às tabelas que estão fisicamente armazenadas no DW, para se conseguir construir as respectivas dimensões e cubos dos modelos desenvolvidos. Ainda antes de se partir para a ligação ao DW e extrair as tabelas necessárias para os modelos a definir, a seguinte figura mostra a organização de uma solução em Microsoft *Analysis Services*:

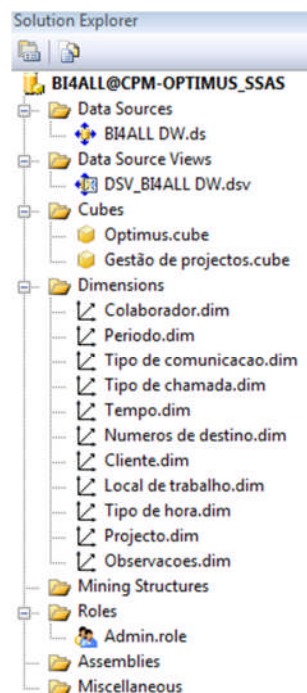


Figura 46 - Explorador da solução *analysis services* do caso de estudo

Tem-se então como primeira camada a ligação ao DW, que na imagem acima está presente em “*Data Sources*”, sendo a ligação efectuada no BI4ALL DW.ds. Depois dessa ligação, são escolhidas as tabelas contidas no DW para construir o modelo pretendido – isso é realizado em “*Data Source Views*” (DSV) – DSV_BI4ALL DW.dsv. De seguida, em “*Dimensions*” são criadas as dimensões necessárias para o modelo, tendo como base o que foi carregado no DSV. Por último, cria-se os cubos em “*Cubes*”, onde se selecciona as factuais pretendidas e que estão ligadas às dimensões segundo o DSV definido anteriormente. Todos estes aspectos vão ser referidos com recurso a figuras já de seguida.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Portanto, o primeiro passo é fazer a ligação ao DW em “Data Sources” e posteriormente, criar o DSV o qual vai ter todas as tabelas necessárias para os nossos modelos. Em termos de exposição, tem-se o modelo global ilustrado na seguinte figura:

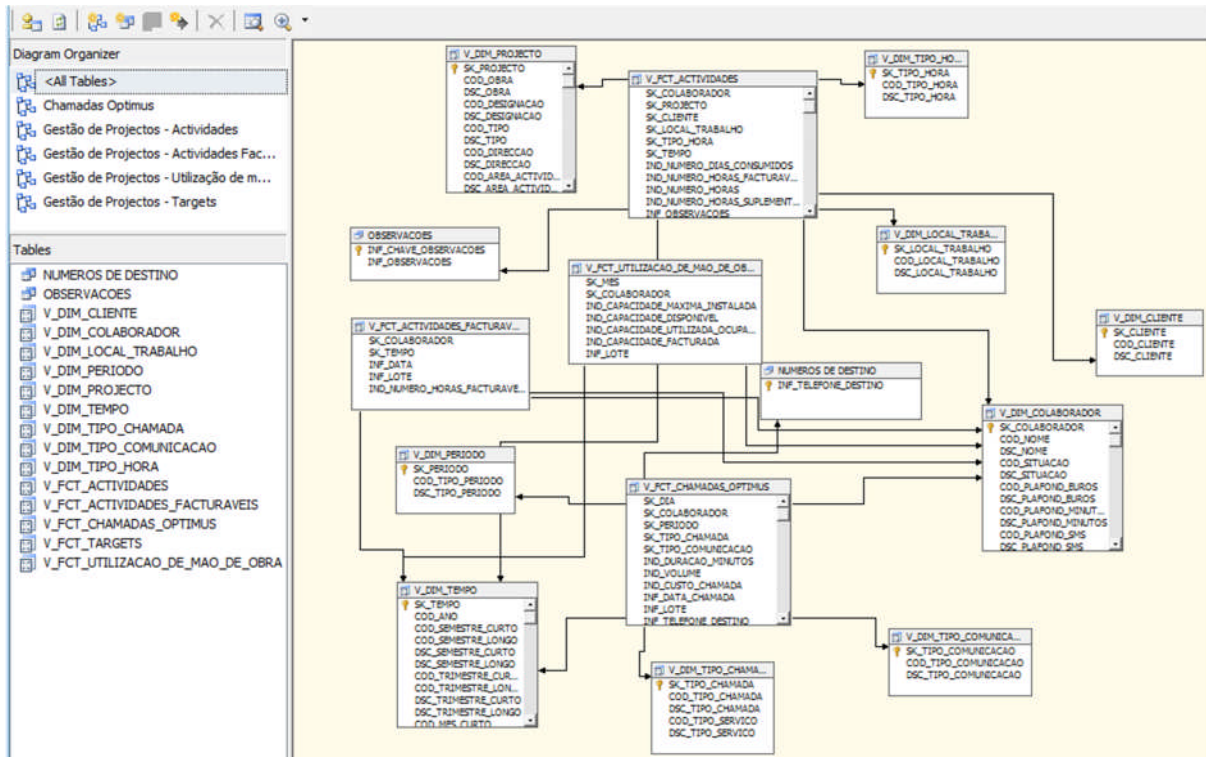


Figura 47 - Data source view dos modelos

Tem-se então do lado esquerdo o “Diagram Organizer”, onde estão todos os modelos definidos para este caso de estudo e as respectivas tabelas necessárias para a definição dos mesmos. No centro tem-se a ligação entre todas as factuais com as suas respectivas dimensões. Estes diagramas servem depois como fonte para a construção das dimensões, ilustrado na figura abaixo com a dimensão tempo como exemplo:

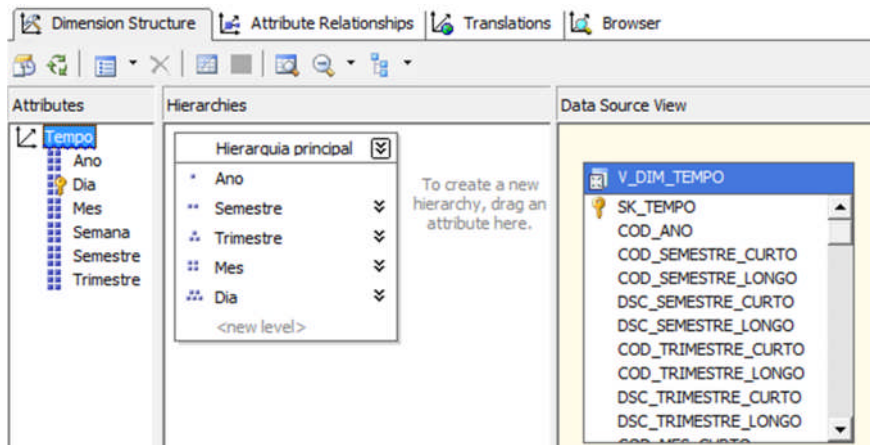


Figura 48 - Estrutura da dimensão tempo

Tem-se então na construção da dimensão onde, no lado direito, se encontra a tabela contida no DSV, com todos os campos disponíveis e que está fisicamente alojado no DW. Mais à esquerda tem-se os atributos necessários para os utilizadores finais poderem dar contexto aos dados. Esses atributos possuem uma designação mais apropriada para esses utilizadores e são definidos com base na seguinte lógica: o atributo Ano tem como chave o atributo COD_ANO contido no DSV, sendo o valor apresentado quando consultado no cubo o que está contido no DSC_ANO no DSV.

As hierarquias também são aqui definidas, tendo-se a seguinte imagem como exemplo para a hierarquia presente na dimensão tempo:

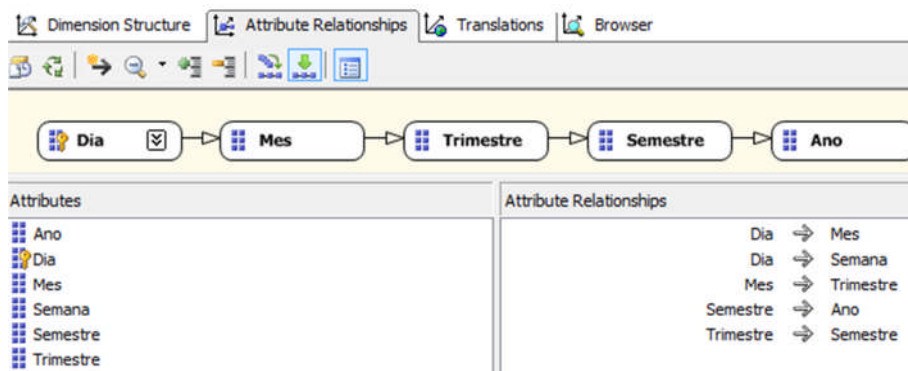


Figura 49 - Hierarquia da dimensão tempo

Depois das dimensões criadas, cria-se o cubo para o(s) modelo(s) em questão, ilustrado abaixo com o cubo de gestão de projectos:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

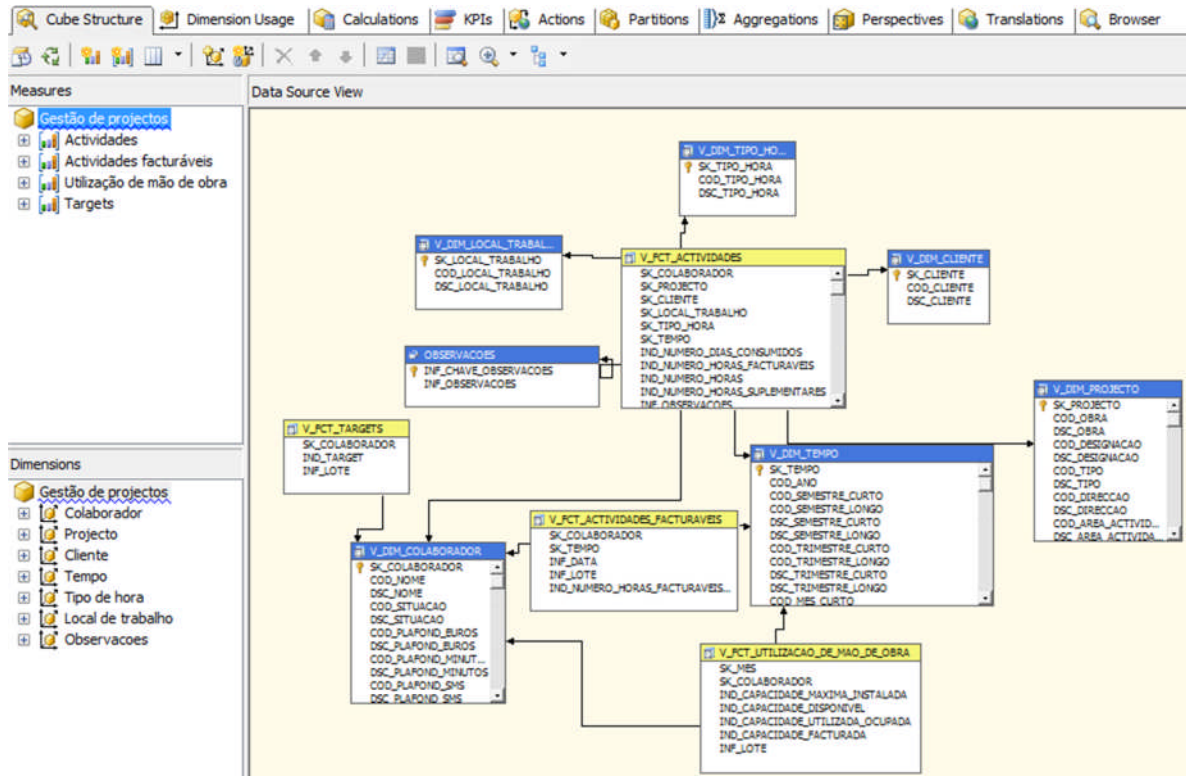


Ilustração 50 - Estrutura do cubo do modelo de gestão de projectos

Tem-se então a estrutura do cubo, com as métricas organizadas por sub-modelo do lado esquerdo, e todas as dimensões associadas aos modelos também do lado esquerdo na parte inferior. No centro tem-se os vários modelos contidos no DSV. Das várias funcionalidades que esta tecnologia oferece em termos de construção de cubos e que foi utilizada neste caso de estudo, foi a parte das calculações realizadas no cubo. Estes cálculos são feitos em tempo real quando se consulta o cubo, não sendo armazenados fisicamente. Um exemplo destes cálculos são as métricas derivadas identificadas no levantamento de requisitos e, exemplificadas na imagem abaixo:

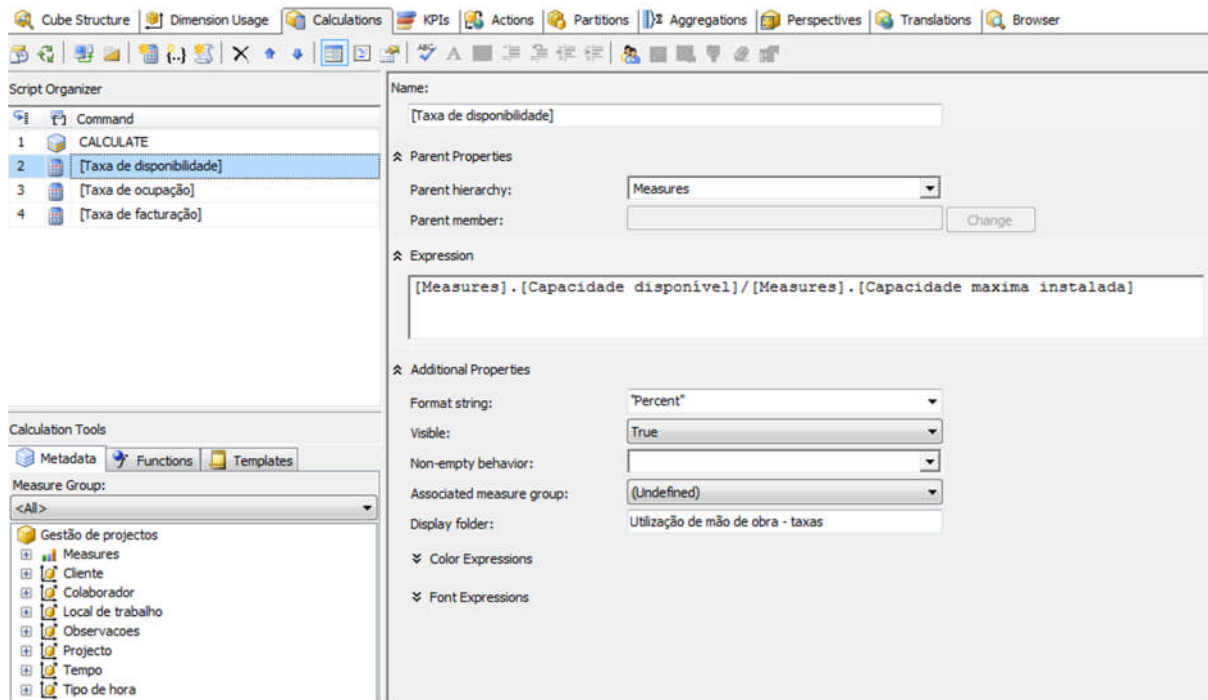


Figura 51 - Cálculos efectuados no cubo do modelo de gestão de projectos

4.7 Designing and Developing Business Intelligence Applications

Este ponto é onde se define como se vai aceder aos dados carregados na fase de ETL e modelados no DW (*design*) e, onde se desenvolve também a solução para o utilizador a usar nesse mesmo acesso ao DW (*developing*). No caso deste trabalho, foi feito o *design* das aplicações em dois momentos, tendo sido feito primeiramente para o modelo de controlo e redução de custos, e depois para a gestão de projectos. Para cada um dos modelos, só se iniciou esta fase, por questões temporais e de teor prático, quando a fase de ETL para os mesmos estava terminada e validada. Sendo assim, para o modelo de controlo e redução de custos esta fase foi feita na sua plenitude, sendo a aplicação desenvolvida já se encontrando em produção e a ser utilizada pelos utilizadores actualmente. Dado ao limite do tempo para a realização desta dissertação, a fase de conclusão do ETL de gestão de projectos foi bastante tardia, influenciando a realização deste ponto na sua totalidade. Mesmo assim, através da solução de SSIS, os utilizadores conseguem através do portal criar os seus relatórios à medida das suas necessidades, e fazer consultas *ad-hoc* à informação contida no DW. Vai-se agora proceder ao detalhe deste ponto sobre os dois modelos.

Para o modelo de controlo e redução de custos, após as reuniões de levantamento para esta fase, foram identificadas três camadas de relatórios:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

- Relatório geral;
- Relatório ao nível do colaborador e,;
- Relatório de detalhe por tipo de da chamada e por colaborador.

O relatório geral, indica para um dado período temporal permite ao utilizador verificar de uma forma geral e resumida, os indicadores mais relevantes para este tipo de análise e que se pode verificar na seguinte figura (dividida em dois para melhor visualização neste documento):



Relatório das chamadas realizadas por colaboradores - Controlo e redução de custos

Relatório Geral - Período em análise: 2011 / 2ºTrim

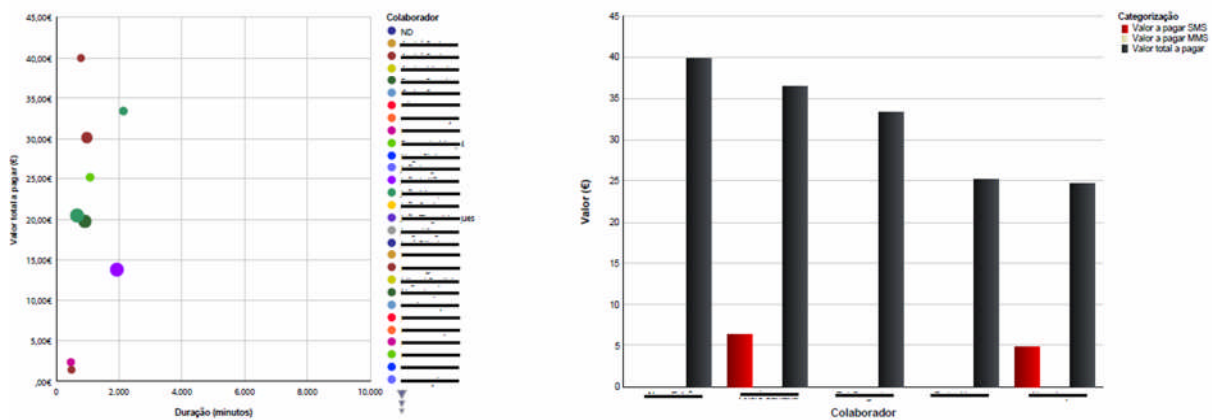


Figura 52 - Exemplo de um relatório geral sobre o modelo de controlo e redução de custos (1/2)

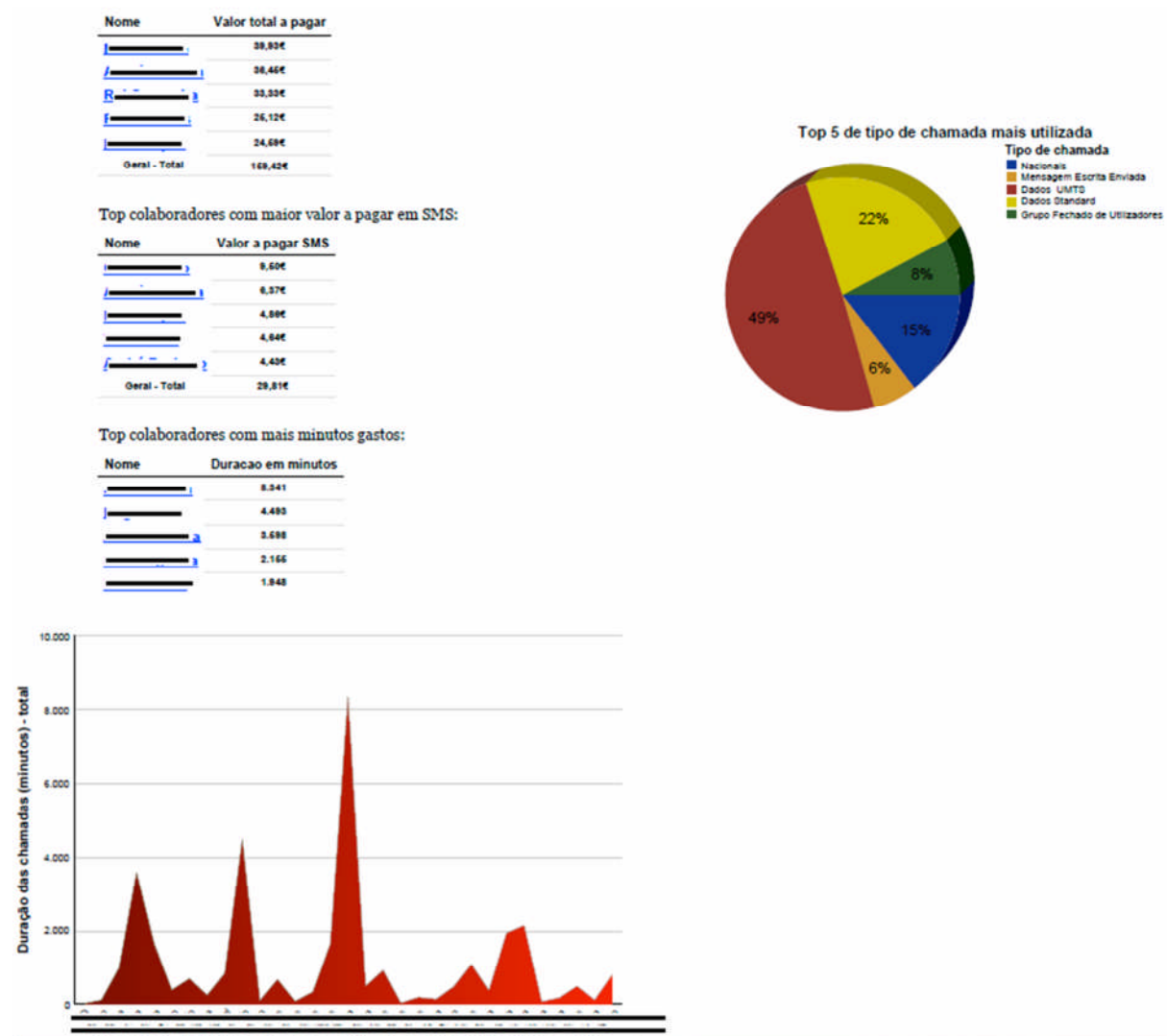


Figura 53 - Exemplo de um relatório geral sobre o modelo de controlo e redução de custos (2/2)

Através deste primeiro relatório, é possível fazer *drill-down* para um colaborador e, analisar para o período temporal escolhido no relatório geral, os indicadores mais relevantes e de uma forma mais detalhada para análise e rastreabilidade. Um exemplo deste relatório encontra-se nas seguintes imagens abaixo (dividida em dois para melhor visualização neste documento):

em formato excel (para uma melhor facilidade de manuseamento da informação), encontra-se ilustrado na seguinte figura:

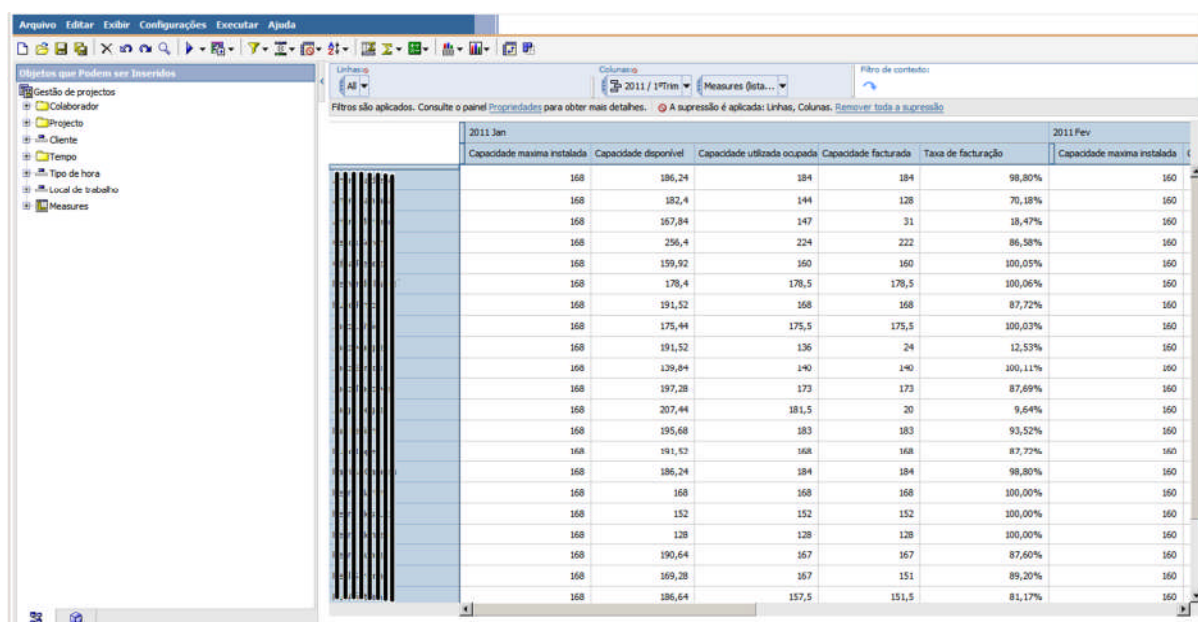
| | A | B | C |
|----|---|---|-----------|
| 1 | Relatório ao nível da chamada | | |
| 2 | Controlo e redução de custos | | |
| 10 | Colaborador: Mendes Lopes | | |
| 12 | Análise do período: All | | |
| 14 | Duração em minutos - por tipo de chamada seleccionada | | Nacionais |
| 15 | 2009-12-02 00:00:00.000 | | 33 |
| 16 | | | 3 |
| 17 | | | 3 |
| 18 | | | 1 |
| 19 | 2009-12-03 00:00:00.000 | | 5 |
| 20 | | | 2 |
| 21 | | | 9 |
| 22 | | | 3 |
| 23 | | | 10 |
| 24 | | | 1 |
| 25 | 2009-12-04 00:00:00.000 | | 1 |
| 26 | 2009-12-06 00:00:00.000 | | 2 |
| 27 | 2009-12-07 00:00:00.000 | | 7 |
| 28 | 2009-12-09 00:00:00.000 | | 2 |
| 29 | | | 8 |
| 30 | | | 3 |
| 31 | | | 4 |
| 32 | | | 2 |
| 33 | 2009-12-10 00:00:00.000 | | 10 |
| 34 | | | 3 |
| 35 | | | 1 |
| 36 | 2009-12-11 00:00:00.000 | | 2 |
| 37 | | | 7 |

Figura 56 – Exemplo de um relatório ao nível da chamada de um colaborador

É possível através da solução desenvolvida no *software* Cognos exportar para formatos PDF, WORD, EXCEL para um melhor manuseamento da informação.

Para a gestão de projectos, com a nova solução implementada é apenas possível ainda efectuar *queries ad-hoc*. Não foi possível re-fazer os relatórios utilizados pelo sistema antigo devido ao pouco tempo disponível para validar se esses relatórios ainda eram utilizados e, se era necessário melhorar algum aspecto, etc. Não fazia sentido avançar para este passo sem primeiro fazer um levantamento exaustivo em termos de *design* da solução. Mesmo assim, é possível a qualquer utilizador abrir o cubo através do *Cognos Analysis studio*, e navegar nos dados realizando as suas análises e exportando os dados para vários tipos de formato. Isto encontra-se ilustrado na seguinte figura:

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI



The screenshot displays the IBM Cognos Analysis Studio interface. On the left, there is a navigation pane with a tree view containing folders like 'Gestão de projectos', 'Colaborador', 'Projecto', 'Cliente', 'Tempo', 'Tipo de hora', 'Local de trabalho', and 'Measures'. The main area shows a data table with the following columns: '2011 Jan', 'Capacidade máxima instalada', 'Capacidade disponível', 'Capacidade utilizada ocupada', 'Capacidade facturada', 'Taxa de facturação', and '2011 Fev'. The table contains 18 rows of data. The interface also includes a menu bar at the top with options like 'Arquivo', 'Editar', 'Exibir', 'Configurações', 'Executar', and 'Ajuda'. A status bar at the bottom indicates 'Filtros são aplicados. Consulte o painel Propriedades para obter mais detalhes.' and 'A supressão é aplicada: Linhas, Colunas. Remover toda a supressão'.

| 2011 Jan | Capacidade máxima instalada | Capacidade disponível | Capacidade utilizada ocupada | Capacidade facturada | Taxa de facturação | 2011 Fev |
|----------|-----------------------------|-----------------------|------------------------------|----------------------|--------------------|----------|
| | 168 | 186,24 | 184 | 184 | 98,80% | 160 |
| | 168 | 182,4 | 144 | 128 | 70,18% | 160 |
| | 168 | 167,84 | 147 | 31 | 18,47% | 160 |
| | 168 | 256,4 | 224 | 222 | 86,58% | 160 |
| | 168 | 159,92 | 160 | 160 | 100,05% | 160 |
| | 168 | 178,4 | 178,5 | 178,5 | 100,06% | 160 |
| | 168 | 191,52 | 168 | 168 | 87,72% | 160 |
| | 168 | 175,44 | 175,5 | 175,5 | 100,03% | 160 |
| | 168 | 191,52 | 136 | 24 | 12,53% | 160 |
| | 168 | 129,84 | 140 | 140 | 100,11% | 160 |
| | 168 | 197,28 | 173 | 173 | 87,69% | 160 |
| | 168 | 207,44 | 181,5 | 20 | 9,64% | 160 |
| | 168 | 195,68 | 183 | 183 | 93,52% | 160 |
| | 168 | 191,52 | 168 | 168 | 87,72% | 160 |
| | 168 | 186,24 | 184 | 184 | 98,80% | 160 |
| | 168 | 168 | 168 | 168 | 100,00% | 160 |
| | 168 | 152 | 152 | 152 | 100,00% | 160 |
| | 168 | 128 | 128 | 128 | 100,00% | 160 |
| | 168 | 190,64 | 167 | 167 | 87,60% | 160 |
| | 168 | 169,28 | 167 | 151 | 89,20% | 160 |
| | 168 | 186,64 | 157,5 | 151,5 | 81,17% | 160 |

Figura 57 - Navegação sobre os dados do modelo de gestão de projectos usando o IBM Cognos Analysis Studio

4.8 Considerações finais sobre a metodologia adoptada

Como foi possível constatar ao longo deste documento, para cada passo da metodologia de Kimball et al. (2008), foram realçados os pontos mais importantes para este trabalho. Para as restantes componentes da metodologia não mencionados, isso foi devido a não fazer sentido realçar o *Growth* (crescimento) e *Maintenance* (manutenção). O primeiro porque crescimento não se aplica pois não foi nada evoluído, mas sim re-feito de raiz. O segundo porque apenas alguns componentes ainda só nesta fase de submissão da dissertação é que passaram a produção, sendo a sua integração ainda não se encontrando completa a 100% em termos de relatórios e modelos disponibilizados.

5 Conclusões

Esta metodologia comprovou ser adequada, pois permitiu estruturar o trabalho com mais precisão, o que aliado com a dedicação dos patrocinadores de negócio da BI4ALL e o orientador da dissertação no ISCTE-IUL, resultou num trabalho bem sucedido. A metodologia de Kimball *et al.* (2008) sendo geral e abrangente, permite ser flexível para os vários tipos de projectos de BI no mercado, tal como foi possível constatar com este trabalho. Não obstante a metodologia deveria contemplar uma fase terminal para validação da solução implementada, para determinar com uma maior clareza uma avaliação dos resultados obtidos das várias iterações pelo ciclo de vida da metodologia de Kimball *et al.* (2008).

De acordo com os objectivos propostos, obteve-se o seguinte resultado:

Foram identificados, modelados e desenvolvidos os modelos dimensionais que permitiram dar à organização uma nova capacidade de monitorização e análise nas áreas de controlo e redução de custos e, gestão de projectos, resultando assim num sistema de BI *end-to-end*, desde a extração dos dados dos sistemas fonte, até à camada de apresentação onde se encontram os relatórios e o acesso aos dados via *ad-hoc*, como todos os KPIs pedidos pelo negócio contidos nos respectivos modelos.

Como vantagens de todo este trabalho, foi pedido *feedback* a alguns utilizadores chave que passaram directamente a usufruir no dia-a-dia com o desenvolvimento efectuado:

"(...) O Projecto de controlo e redução de custos de chamadas telefónicas que o Rui implementou na BI4ALL, levou a que a empresa conseguisse apurar os custos reais das chamadas por colaborador, e dessa forma conseguisse melhorar a sua eficácia na distribuição dos valores de acordo com o negociado com a empresa. Na solução interna de gestão de projectos, solução crítica para o nosso negócio, o Rui refez a solução existente melhorando dessa forma a capacidade analítica da BI4ALL no controle da gestão dos projectos. Para terem uma ideia da importância desta solução agora refeita pelo Rui, é a partir daqui que saem todos os relatórios para a contabilidade de gestão de horas, bem como os mapas de controlo de horas para os clientes finais. No final ficamos com uma solução que responde a uma série de necessidades internas para responder aos pedidos de esclarecimento dos clientes todos os meses." – José Oliveira, Director-Geral da BI4ALL.

“Esta implementação veio facilitar em muito o nosso trabalho dado que neste momento conseguimos ter rápido acesso à informação, obter relatórios totais e apenas de um colaborador,

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

bem como uma maior exactidão no controlo de custos efectuados pelos nossos colaboradores.” – Dora Palma, acessora de comunicação da BI4ALL.

“Os pontos positivos são um melhor controlo da factura da Optimus, o facto de conseguirmos ter a percepção real de quem utiliza e como são utilizados/distribuídos os nossos minutos/sms em comunicações, o que se está a concretizar na optimização do processo de cobrança de plafonds extra aos Colaboradores e numa melhor negociação com os operadores de comunicações.” – *Partner* da BI4ALL.

Através destes testemunhos, é de realçar o contributo positivo que esta dissertação teve directamente para a empresa BI4ALL, onde já é extraído valor de todos os desenvolvimentos efectuados. Comprovou-se também que a metodologia de Kimball *et al.* (2008) em termos práticos, se mostrou vantajosa em termos de gestão de tarefas desde o arranque deste trabalho, mostrando-se flexível para projectos de BI desta natureza.

Não só a BI4ALL retirou valor deste trabalho, como o centro ADETTI-IUL poderá agora aproveitar todo o conhecimento que aqui se encontra e, aplicá-lo de acordo com as suas actuais necessidades, retirando assim mais valor do seu negócio.

6 Recomendações

Com este trabalho, foi possível constatar que na BI4ALL existem ainda mais áreas funcionais que necessitam de ser modeladas, seguindo assim a linha de desenvolvimento adoptada para as áreas de controlo e redução de custos e, gestão de projectos. Nomeadamente, as áreas comercial e da facturação, têm assim uma base de desenvolvimento estável e passível de evolução.

Esta possibilidade de acrescentar novas áreas de modelação ao desenvolvimento já realizado com este trabalho, prova que esta dissertação contribuiu com uma solução flexível para qualquer desenvolvimento adicional que já tenha sido feito até ao momento na BI4ALL.

De referir ainda que este trabalho não só será um instrumento útil para quaisquer novos desenvolvimentos na BI4ALL, como também para qualquer novo desenvolvimento noutra instituição com necessidades distintas da BI4ALL.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Bibliografia

(Adamson 2010) Adamson, C.; *Star Schema – The Complete Reference*, McGrawHill, 2010.

(AMR 2006) AMR, *Workforce management landscape: The right people in the right place at the right time*, AMR Res. Rep., 2006.

(Andonov-Acev *et al.* 2008) Andonov-Acev, D.; Buckovska, A.; Blagojevic, Z.; Kraljevski, V.; *Enterprise performance monitoring*, Proceedings of the ITI, 2008.

(BI-INSIDER.COM 2001) BI-INSIDER.COM, *Dimensional Modeling and Data Warehouses*, 2011. Obtido em <http://bi-insider.com/posts/dimensional-modeling-and-data-warehouses/>.

(Bititci 1997) Bititci, U.; *Integrated performance measurement systems: a development guide*, International Journal of Operations, 1997.

(Bonifati 2001) Bonifati, A.; *Designing Data Marts for Data Warehouses*, ACM Transactions on Software Engineering and Methodology, 2001.

(Bourne *et al.* 2003) Bourne, M.; Franco, M.; e Wilkes, J.; *Corporate Performance Management*, Meas Bus Excell, 7: 15-21, 2003.

(Caldeira 2008) Caldeira, C.; *Data Warehousing - conceitos e modelos*, Edições Sílabo, 2008.

(Cardoso 2008) Cardoso, E.; *Arquitecturas de DW*, Slides apresentados nas aulas do mestrado de engenharia informática, 2008.

(Cardoso 2009)¹ Cardoso, E.; *Dimensional Modelling – part 4*, Slides apresentados nas aulas do mestrado de engenharia informática, 2009.

(Cardoso 2009)² Cardoso, E.; *Introduction to Business Intelligence*, Slides apresentados nas aulas do mestrado de engenharia informática, 2009.

(Cardoso 2009)³ Cardoso, E.; *Introduction to Dimensional Modeling*, Slides apresentados nas aulas do mestrado de engenharia informática, 2009.

(Che *et al.* 2009) Chee, T.; Chan, L.; Chuah, M.; Tan, C.; Wong, S.; Yeoh, W.; *Business Intelligence Systems: State-Of-The-Art Review And Contemporary Applications*, Symposium on Progress in Information & Communication Technology, 2009.

(Chen *et al.* 2005) Chen B.; Tai R.; Harrison R.; e Pan Y.; *Novel hybrid hierarchical-k-means clustering method (h-k-means) for microarray analysis*, IEEE BCSBW, 2005.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

(Cody *et al.* 2002) Cody, W.; Kreulen, J.; Krishna, V.; e Spangler, W.; *The Integration of business intelligence and knowledge management*, IBM Systems Journal, 2002.

(Cognos 2008) Cognos (organization), *The evolution of the CPM system*, publicado em <http://www.faktos.com/shared/downloads/evolution%20of%20cpm%20system.pdf>, 2008.

(Dransfield 2000) Dransfield, R.; *Human resource management*, Heinemann, 2000.

(Golfarelli *et al.* 2004) Golfarelli, M.; Rizzi, S.; e Cella, I.; *Beyond Data Warehousing: What's Next in Business Intelligence?*, DEIS - University of Bologna, 2004.

(Heß 2006) Heß, H.; *From corporate strategy to process performance – what comes after business intelligence?*, IDS Scheer AG, 2006.

(Hu, Lu 2010) Hu, J.; e Lu, Y.; *Performance management of IT services delivery*, IBM Thomas J. Watson Research Center, 2010.

(Charlesworth *et al.* 2004) Charlesworth, I.; Edwards, R.; Kellet, A.; *Corporate Performance Management - A new approach to business control and planning*, SAS Events, 2004.

(IBM 2010) IBM, *IBM Cognos Business Intelligence – Guia de Introdução*, Versão 10.0.1, IBM, 2010.

(Inmon 1996) Inmon, W.; *The Data Warehouse and Data Mining*, Communications of the ACM, 1996.

(Jingliang *et al.* 2010) Jingliang, R.; Xingzheng, X.; e Chenchun, L.; *A Research on the Strategic Performance Management Systems from the Perspective of Core Competencies*, Management and Service Science (MASS), 2010 International Conference, 2010.

(Kang, Han 2008) Kang, J.; e Han, K.; *A Business Activity Monitoring System Supporting Real-Time Business Performance Management*, 2008

(Kimball *et al.* 2008) Kimball, R.; Ross, M.; Thornthwaite, W.; Mundy, J.; Becker, B.; *The Data Warehouse Lifecycle Toolkit*, 2nd Edition, 2008.

(Larson 2009) Larson, B.; *Delivering Business Intelligence with Microsoft SQL Server 2008*, McGraw-Hill, 2009.

(Lawton 2006) Lawton, G.; *Making Business Intelligence More Useful*, Industry Trends, 2006.

(Lukawiecki 2009) Lukawiecki, R.; *Soluções Microsoft para Business Intelligence e Gestão de Desempenho – dos dados às decisões*, Microsoft BI Event, Tagus Park, 2009.

Bibliografia

(Melchert 2004) Melchert, F.; *Aligning Process Automation and Business Intelligence to Support Corporate Performance Management*, Proceedings of the Tenth Americas Conference on Information Systems, 2004.

(Microsoft TechNet 2011) Microsoft TechNet, Processing Options and Settings, 2011. Obtido em <http://technet.microsoft.com/en-us/library/ms174774.aspx>

(Mosimann *et al.* 2007) Mosimann, R.; Mosimann, P.; e Dussault, M.; *The performance manager – Proven strategies for turning information into higher business performance*, Cognos, 2007.

(Negash 2004) Negash, S.; *Business Intelligence*, Communications of the Association for Information Systems, 2004.

(Oliveira *et al.* 2006) Oliveira, J.; Baptista, M.; e Sezões, C.; *Business Intelligence*, Sociedade Portuguesa de Inovação, 2006.

(Quinn 2008) Quinn, K.; *How Business Intelligence Should Work. The connection between Strategic, Analytical, and Operational Initiatives*, Information Builders White Paper, WebFOCUS, 2008.

(Sallam *et al.* 2011) Sallam, R.; Richardson J.; Hagerty, J.; e Hostmann; *Magic Quadrant for Business Intelligence Plataforms*, Gartner RAS Core Research Note, 2011.

(Van den Heuvel *et al.* 2001) Van den Heuvel, W.; Hasselbring, W.; e Kumar, K.; *Enterprise Application Integration: Introduction to the Minitrack*, IEEE, 2001.

(Watson, Wixom 2007) Watson, H.; e Wixom, B.; *The Current State of Business Intelligence*, IT Systems Perspectives, 2007.

(White 2006) White, C.; *The Next Generation of Business Intelligence: Operational BI*, BI research, Sybase, 2006.

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

Anexos

Anexos

Anexo A

Hierarquias

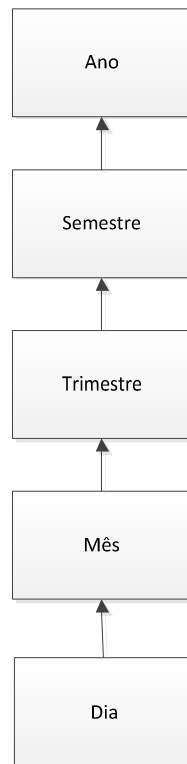


Figura 58 – Hierarquia da dimensão Tempo

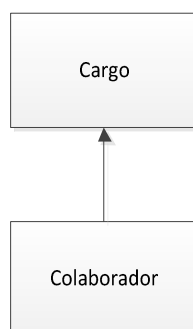


Figura 59 - Hierarquia da dimensão tipo de comunicação

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI



Figura 60 - Hierarquia da dimensão tipo de chamada

ANEXO B

Metadata (DW)

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|------------------------|---|----------|----------------------------|
| SK_DIA | Chave estrangeira da dimensão tempo | INT | 1 |
| SK_COLABORADOR | Chave estrangeira da dimensão colaborador | INT | 1 |
| SK_PERIODO | Chave estrangeira da dimensão período | INT | 1 |
| SK_TIPO_CHAMADA | Chave estrangeira da dimensão tipo de chamada | INT | 1 |
| SK_TIPO_COMUNICACAO | Chave estrangeira da dimensão tipo de comunicação | INT | 1 |
| IND_DURACAO_MINUTOS | Valor da duração da chamada efectuada | INT | 120 |
| IND_VOLUME | Valor do volume de dados consumidos na chamada | FLOAT | 12,34 |
| IND_CUSTO_CHAMADA | Valor do custo associado à chamada realizada depois de o colaborador ter passado o seu <i>plafond</i> . | FLOAT | 1,23 |
| IND_CUSTO_CHAMADA_REAL | Valor real da chamada mesmo antes de ter passado o <i>plafond</i> do colaborador | FLOAT | 45,3 |
| INF_DATA_CHAMADA | Campo informativo indicando qual o dia em que a chamada foi realizada | DATETIME | 2010-09-03 00:00:00.000 |
| INF_LOTE | Campo informativo sobre qual o lote carregado | INT | 65 |
| INF_TELEFONE_DESTINO | Campo informativo indicando qual o número de telefone de destino da chamada realizada | NVARCHAR | 90 000 00 00 |
| INF_TELEFONE_EMISSOR | Campo informativo indicando qual o número de telefone do emissor da chamada realizada | NVARCHAR | 99 999 99 99 |

Tabela 8 - FCT_CHAMADAS_OPTIMUS – factual do modelo de controlo e redução de custos

Anexos

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|--------------------------------|--|----------|-------------------------|
| SK_COLABORADOR | Chave estrangeira da dimensão colaborador | INT | 1 |
| SK_PROJECTO | Chave estrangeira da dimensão projecto | INT | 1 |
| SK_CLIENTE | Chave estrangeira da dimensão cliente | INT | 1 |
| SK_LOCAL_TRABALHO | Chave estrangeira da dimensão local de trabalho | INT | 1 |
| SK_TIPO_HORA | Chave estrangeira da dimensão tipo de hora | INT | 1 |
| SK_TEMPO | Chave estrangeira da dimensão tempo | INT | 1 |
| IND_NUMERO_DIAS_CONSUMIDOS | Número de dias consumidos numa actividade | FLOAT | 1,23 |
| IND_NUMERO_HORAS_FACTURAVEIS | Quantidade de horas que foram facturáveis a clientes numa actividade | FLOAT | 2,5 |
| IND_NUMERO_HORAS | Número de horas gastas numa actividade | FLOAT | 5,5 |
| IND_NUMERO_HORAS_SUPLEMENTARES | Número de horas suplementares numa actividade | INT | 2 |
| INF_OBSERVACOES | Campo informativo indicando o que foi feito numa dada actividade | VARCHAR | “Instalação de um ERP” |
| INF_DATA | Campo informativo indicando qual o dia em que a actividade foi realizada | DATETIME | 2010-09-03 00:00:00.000 |
| INF_LOTE | Campo informativo sobre qual o lote carregado | INT | 70 |

Tabela 9 - FCT_ACTIVIDADES – factual do modelo de gestão de projectos (sub-modelo de Actividades)

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|----------------------------------|---|----------|-------------------------|
| SK_COLABORADOR | Chave estrangeira da dimensão colaborador | INT | 1 |
| SK_TEMPO | Chave estrangeira da dimensão tempo | INT | 1 |
| IND_NUMERO_HORAS_FACTURAVEIS_DIA | Quantidade de horas que foram facturáveis a clientes numa actividade facturável | FLOAT | 3 |
| INF_DATA | Campo informativo indicando qual o dia em que a actividade foi realizada | DATETIME | 2010-09-03 00:00:00.000 |
| INF_LOTE | Campo informativo sobre qual o lote carregado | INT | 70 |

Tabela 10 - FCT_ACTIVIDADES_FACTURAVEIS – factual do modelo de gestão de projectos (sub-modelo de Actividades Facturáveis)

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|----------------|---|-------|------|
| SK_COLABORADOR | Chave estrangeira da dimensão colaborador | INT | 1 |
| IND_TARGET | Quantido de horas que foram facturáveis a clientes numa actividade facturável | FLOAT | 3 |
| INF_LOTE | Campo informativo sobre qual o lote carregado | INT | 70 |

Tabela 11 - FCT_TARGETS – factual do modelo de gestão de projectos (sub-modelo de Targets)

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|----------------------------------|---|------|------|
| SK_MES | Chave estrangeira da dimensão tempo | INT | 1 |
| SK_COLABORADOR | Chave estrangeira da dimensão colaborador | INT | 1 |
| IND_CAPACIDADE_MAXIMA_INSTALADA | Quantidade maxima de mão-de-obra disponível pela BI4ALL | INT | 450 |
| IND_CAPACIDADE_DISPONIVEL | Quantidade disponível de mão-de-obra pela BI4ALL | INT | 320 |
| IND_CAPACIDADE_UTILIZADA_OCUPADA | Quantidade ocupada/utilizada de mão-de-obra pela BI4ALL | INT | 250 |
| IND_CAPACIDADE_FACTURADA | Quantidade de mão-de-obra que é facturável pela BI4ALL | INT | 150 |

Tabela 12 - FCT_UTILIZACAO_DE_MAO_DE_OBRA – factual do modelo de gestão de projectos (sub-modelo de Utilização de mão-de-obra)

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|-------------|------------------------------------|----------|---------|
| SK_CLIENTE | <i>Chave substituta</i> do cliente | INT | 1 |
| COD_CLIENTE | Código do nome do cliente | INT | 1 |
| DSC_CLIENTE | Descrição do nome do cliente | NVARCHAR | “ISCTE” |

Tabela 13 - DIM_CLIENTE – dimensão do modelo de gestão de projectos

Anexos

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|---------------------|--|-------------|---------------------------|
| SK_COLABORADOR | Chave substituta do colaborador | INT | 1 |
| COD_NOME | Código do nome do colaborador | INT | 1 |
| DSC_NOME | Descrição do nome do colaborador | NVARCHAR | “Rui Sousa” |
| COD_DEPARTAMENTO | Código do departamento onde se encontra alocado o colaborador | INT | 1 |
| DSC_DEPARTAMENTO | Descrição do departamento onde se encontra alocado o colaborador | NVARCHAR | “Business Intelligence” |
| COD_SITUACAO | Código da situação actual do colaborador | NVARCHAR | “Activo” |
| DSC_SITUACAO | Descrição da situação actual do colaborador | NVARCHAR | “Situação actual: Activo” |
| COD_PLAFOND_EUROS | Código do plafond em euros | INT | 0 |
| DSC_PLAFOND_EUROS | Descrição do plafond em euros | NVARCHAR | “Ilimitado” |
| COD_PLAFOND_MINUTOS | Código do plafond em minutos | INT | 0 |
| DSC_PLAFOND_MINUTOS | Descrição do plafond de minutos disponíveis | NVARCHAR | “Ilimitado” |
| COD_PLAFOND_SMS | Código do plafond das mensagens escritas | INT | 0 |
| DSC_PLAFOND_SMS | Descrição do plafond das mensagens escritas disponíveis | NVARCHAR | “Ilimitado” |
| COD_USERNAME | Código do username | NVARCHAR | “rsousa” |
| DSC_USERNAME | Descrição do username | NVARCHAR | “rsousa” |
| COD_TELEFONE | Código do telefone | INT | 99 999 99 99 |
| DSC_TELEFONE | Descrição do telefone | NVARCHAR | “99 999 99 99” |
| COD_CARGO | Código do cargo do colaborador | INT | 1 |
| DSC_CARGO | Descrição do cargo do colaborador | NVARCHAR | Consultor |

Tabela 14 - DIM_COLABORADOR – dimensão do modelo de gestão de projectos e do controlo e redução de custos

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|--------------------|---------------------------------------|-------------|-------------|
| SK_LOCAL_TRABALHO | Chave substituta do local de trabalho | INT | 1 |
| COD_LOCAL_TRABALHO | Código do local de trabalho | NVARCHAR | “*” |
| DSC_LOCAL_TRABALHO | Descrição do local de trabalho | NVARCHAR | “Outro” |

Tabela 15 - DIM_LOCAL_TRABALHO – dimensão do modelo de gestão de projectos

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|------------------|-----------------------------|----------|-----------|
| SK_PERIODO | Chave substituta do periodo | INT | 1 |
| COD_TIPO_PERIODO | Código do periodo | INT | 1 |
| DSC_TIPO_PERIODO | Descrição do periodo | NVARCHAR | “Feriado” |

Tabela 16 - DIM_PERIODO– dimensão do modelo de controlo e redução de custos

Tabela DIM_PROJECTO– dimensão do modelo de controlo e redução de custos

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|---------------------|---|----------|--|
| SK_PROJECTO | Chave substituta do projecto | INT | 207 |
| COD_OBRA | Código de obra do projecto | NVARCHAR | “PRJ247” |
| DSC_OBRA | Descrição da obra do projecto | NVARCHAR | “PRJ247” |
| COD_DESIGNACAO | Código da designação do projecto | NVARCHAR | “BI4ALL - Projecto CPM@BI4ALL v2” |
| DSC_DESIGNACAO | Descrição da designação do projecto | NVARCHAR | “BI4ALL - Projecto CPM@BI4ALL v2” |
| COD_TIPO | Código do tipo de projecto | NVARCHAR | “Estágios” |
| DSC_TIPO | Descrição do tipo de projecto | NVARCHAR | “Estágios” |
| COD_DIRECCAO | Código da direção do projecto | NVARCHAR | “Não Facturável” |
| DSC_DIRECCAO | Descrição da direção do projecto | NVARCHAR | “Não Facturável” |
| COD_AREA_ACTIVIDADE | Código da área de actividade do projecto | INT | 199 |
| DSC_AREA_ACTIVIDADE | Descrição da área de actividade do projecto | NVARCHAR | “Gastos Gerais” |
| COD_SITUACAO | Código da situação do projecto | NVARCHAR | “Em Curso” |
| DSC_SITUACAO | Descrição da situação do projecto | NVARCHAR | “Em Curso” |

Tabela 17 - DIM_PROJECTO– dimensão do modelo de controlo e redução de custos

Anexos

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|---------------------|----------------------------------|----------|---------------------|
| SK_TEMPO | <i>Chave substituta</i> do tempo | INT | 1 |
| COD_SEMANA_LONGO | Código da semana | NVARCHAR | “2005W01” |
| DSC_SEMANA_LONGO | Descrição da semana | NVARCHAR | “2005 / 01ª Semana” |
| COD_MES_LONGO | Código do mês | NVARCHAR | “2005M01” |
| DSC_MES_LONGO | Descrição do mês | NVARCHAR | “2005 Jan” |
| COD_TRIMESTRE_LONGO | Código do trimestre | NVARCHAR | “2005Q1” |
| DSC_TRIMESTRE_LONGO | Descrição do trimestre | NVARCHAR | “2005 / 1ºTrim” |
| COD_SEMESTRE_LONGO | Código do semestre | NVARCHAR | “2005H1” |
| DSC_SEMESTRE_LONGO | Descrição do semestre | NVARCHAR | “2005 / 1ºSem” |
| COD_ANO | Código do ano | NVARCHAR | “2005” |

Tabela 18 - DIM_TEMPO– dimensão do modelo de gestão de projectos e do controlo e redução de custos

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|------------------|--|----------|-----------------------|
| SK_TIPO_CHAMADA | <i>Chave substituta</i> do tipo de chamada | INT | 1 |
| COD_TIPO_CHAMADA | Código do tipo de chamada | NVARCHAR | “Chamadas Efectuadas” |
| DSC_TIPO_CHAMADA | Descrição do tipo de chamada | NVARCHAR | “Chamadas Efectuadas” |
| COD_TIPO_SERVICO | Código do tipo de serviço | NVARCHAR | “Roaming” |
| DSC_TIPO_SERVICO | Descrição do tipo de serviço | NVARCHAR | “Roaming” |

Tabela 19 - DIM_TIPO_CHAMADA– dimensão do modelo de controlo e redução de custos

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|----------------------|--|----------|--------------------|
| SK_TIPO_COMUNICACAO | <i>Chave substituta</i> do tipo de comunicação | INT | 1 |
| COD_TIPO_COMUNICACAO | Código do tipo de comunicação | NVARCHAR | “Kanguru 21.6Mbps” |
| DSC_TIPO_COMUNICACAO | Descrição do tipo de comunicação | NVARCHAR | “Kanguru 21.6Mbps” |

Tabela 20 - DIM_TIPO_COMUNICACAO– dimensão do modelo de controlo e redução de custos

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

| CAMPO | DESCRIÇÃO | TIPO | Ex.: |
|---------------|----------------------------------|----------|--------------------|
| SK_TIPO_HORA | Chave substituta do tipo de hora | INT | 1 |
| COD_TIPO_HORA | Código do tipo de hora | NVARCHAR | “Kanguru 21.6Mbps” |
| DSC_TIPO_HORA | Descrição do tipo de hora | NVARCHAR | “Kanguru 21.6Mbps” |

Tabela 21 - DIM_TIPO_HORA– dimensão do modelo de gestão de projectos

Anexo C

Pacotes da extração e transformação da informação dos sistemas fonte para a *staging area*

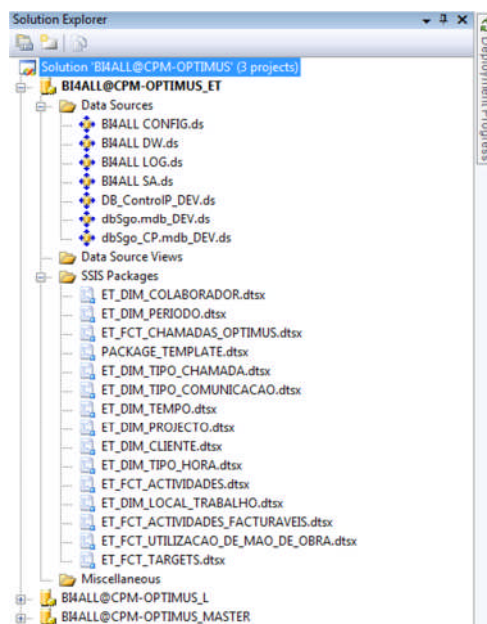


Figura 61 - Pacotes de ET (extração e transformação)

Anexos

Pacotes de carregamento das dimensões e factuais

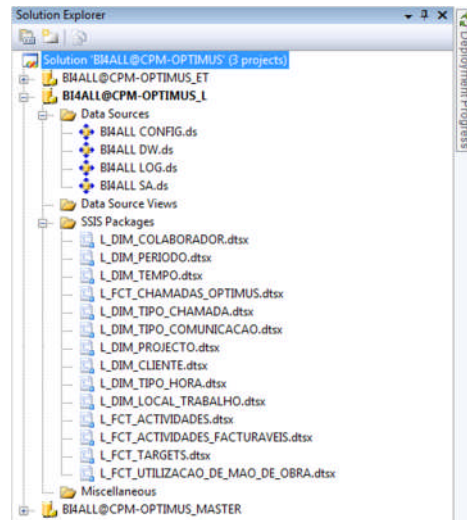


Figura 62 - Pacotes de L (carregamento)

Anexo D

Bases de dados

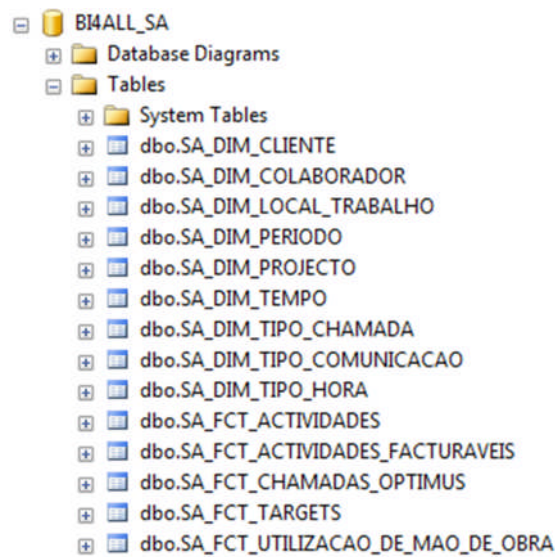


Figura 63 - Base de dados da área de teste (*staging*)

CPM – Aplicação do modelo de Kimball numa empresa de prestação de serviços de BI

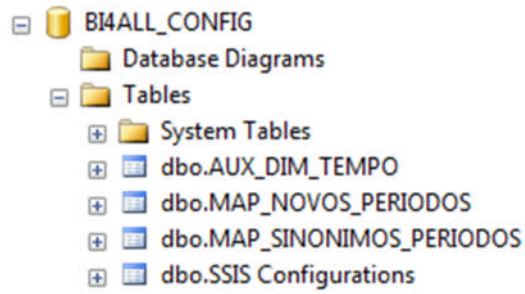


Figura 64 - Base de dados de Config

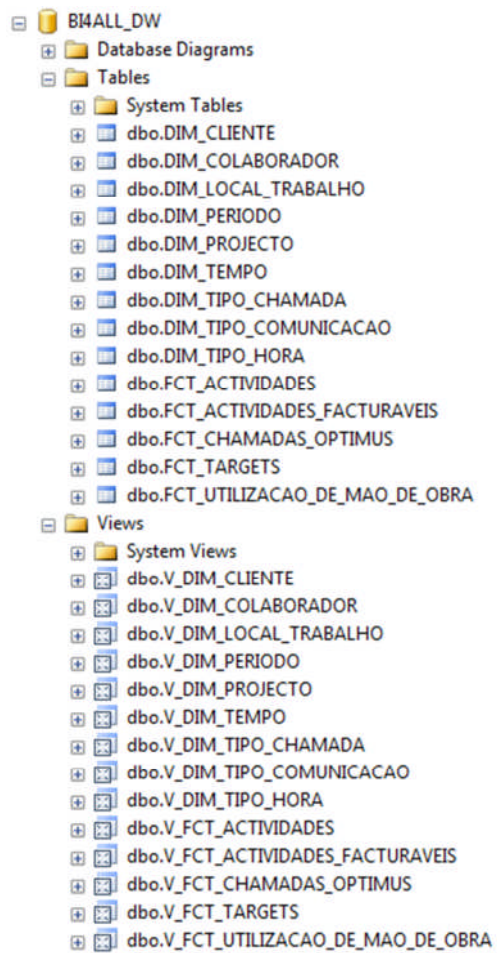


Figura 65 - Base de dados de DW

Anexos

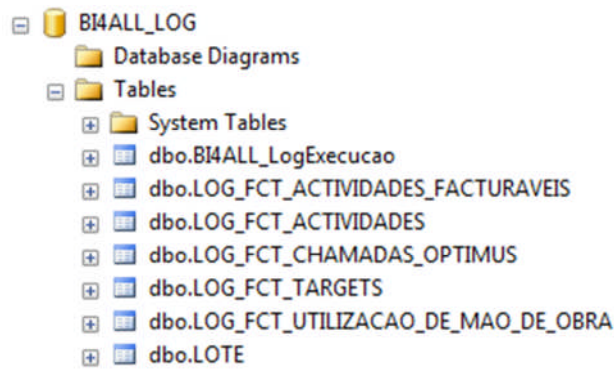


Figura 66 - Base de dados de Log

Anexo E

Exemplo do código de criação de uma *view*

```
CREATE VIEW [dbo].[V_DIM_COLABORADOR]
AS
SELECT SK_COLABORADOR, COD_NOME, DSC_NOME, COD_DEPARTAMENTO,
DSC_DEPARTAMENTO, COD_SITUACAO, DSC_SITUACAO,
COD_PLAFOND_EUROS, DSC_PLAFOND_EUROS, COD_PLAFOND_MINUTOS,
DSC_PLAFOND_MINUTOS, COD_PLAFOND_SMS, DSC_PLAFOND_SMS,
COD_USERNAME, DSC_USERNAME, COD_TELEFONE, DSC_TELEFONE,
COD_CARGO, DSC_CARGO
FROM dbo.DIM_COLABORADOR
UNION
SELECT - 1 AS SK_COLABORADOR, - 1 AS COD_NOME, 'ND' AS
DSC_NOME, - 1 AS COD_DEPARTAMENTO, 'ND' AS DSC_DEPARTAMENTO,
'ND' AS COD_SITUACAO, 'ND' AS DSC_SITUACAO, - 1 AS
COD_PLAFOND_EUROS, 'ND' AS DSC_PLAFOND_EUROS, - 1 AS
COD_PLAFOND_MINUTOS, 'ND' AS DSC_PLAFOND_MINUTOS,
1 AS COD_PLAFOND_SMS, 'ND' AS DSC_PLAFOND_SMS, 'ND' AS
COD_USERNAME, 'ND' AS DSC_USERNAME, - 1 AS COD_TELEFONE, 'ND'
AS DSC_TELEFONE, - 1 AS COD_CARGO, 'ND' AS DSC_CARGO
```