



GRUPO IX

GRUPO DE ESTUDO DE OPERAÇÃO DE SISTEMAS ELÉTRICOS - GOP

BASE DE DADOS HISTÓRICA DA OPERAÇÃO: ANÁLISE DE SOLUÇÕES E PROPOSTAS EVOLUTIVAS

Pedro Daniel Zarur * Luiz Antônio Cordeiro Pereira

CEPEL – Centro de Pesquisas de Energia Elétrica

RESUMO

O avanço tecnológico da supervisão e controle em tempo real, associada à evolução de tecnológica no tratamento da informação, proporciona um cenário com novos desafios e soluções, ideal para a discussão de diretrizes a serem tomadas para o desenvolvimento de sistemas de armazenamento de dados históricos da operação. O objetivo deste trabalho é expor a solução utilizada pelo SAGE - Sistema Aberto de Gerenciamento de Energia, atualmente adotada por várias empresas do setor elétrico brasileiro, e propor uma evolução do produto para uma solução baseada em tecnologia *data warehousing* a fim de atender às necessidades observadas com a experiência no uso da ferramenta.

PALAVRAS-CHAVE

Histórico da Operação, *Data Warehousing*, SCADA, *Data Mining*, SAGE

1.0 - INTRODUÇÃO

Largamente utilizada por diversas empresas do setor, a Base de Dados Histórica da Operação, também denominada BDH, vai muito além de ser apenas uma ferramenta de armazenamento de dados. A BDH constitui-se em uma importante estratégia na integração do departamento de operação com os demais setores da empresa de energia elétrica. Além da função de operação do sistema elétrico, o departamento de operação torna-se um importante provedor de informações para o restante da empresa. As técnicas e ferramentas utilizadas atualmente na implantação de tais bases de dados devem permitir que além de uma visão histórica, seja possível uma vista do processo elétrico próximo ao tempo real. Isso torna a BDH um caminho natural de integração das diversas aplicações computacionais utilizadas pela empresa que utilizam tais informações. A figura 1 mostra um esquema básico dessa integração.

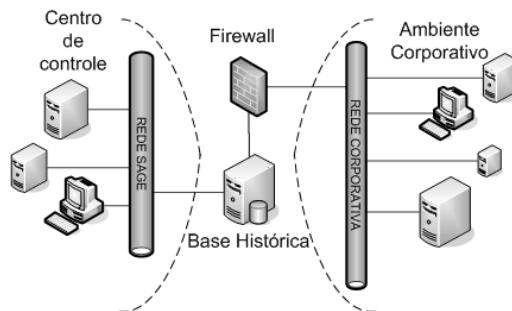


Figura 1: Integração entre o centro de controle e o ambiente corporativo

O objetivo deste trabalho é a relatar da experiência com o desenvolvimento de bancos de dados históricos em diversas empresas do setor que utilizam o SAGE, e propor um caminho evolutivo de uma solução de simples armazenamento de séries temporais em bancos de dados relacionais para um *data warehouse* que disponibiliza informações de forma orientada ao usuário final de forma simples e eficiente a despeito da complexidade e grande volume de dados que envolvem tais bancos de dados.

No próximo capítulo será apresentada a solução atual adotada pelo SAGE para geração de bases históricas. No capítulo seguinte será feita uma breve introdução aos conceitos que envolvem a tecnologia de *data warehousing*. Na terceira parte são mostrados os passos necessários para a implantação de um *data warehouse* com informações históricas da operação a partir da base histórica do SAGE. Finalmente serão apresentadas as conclusões e considerações finais.

2.0 - BDH – BASE DE DADOS HISTÓRICA DO SAGE

A base de dados histórica do SAGE foi desenvolvida com o objetivo de armazenar em bancos de dados relacionais séries temporais representativas da operação do sistema elétrico e sua evolução. Atualmente implantada em várias empresas do setor elétrico brasileiro ela utiliza como repositório de dados um banco de dados relacional de mercado. Atualmente o sistema pode ser integrado aos bancos de dados Oracle, Informix e Postgresql.

2.1 Fluxo da informação do SAGE

A informação no SAGE flui a partir de duas fontes específicas. A primeira é proveniente do carregamento off-line da base de dados fonte do sistema, que contém informações cadastrais sobre a topologia do sistema elétrico. A segunda é proveniente da monitoração do sistema elétrico pelas aplicações do SAGE, ou seja, são informações de tempo real que variam durante o processo de supervisão e controle, de forma on-line. A figura 2 descreve o fluxo da informação do SAGE.

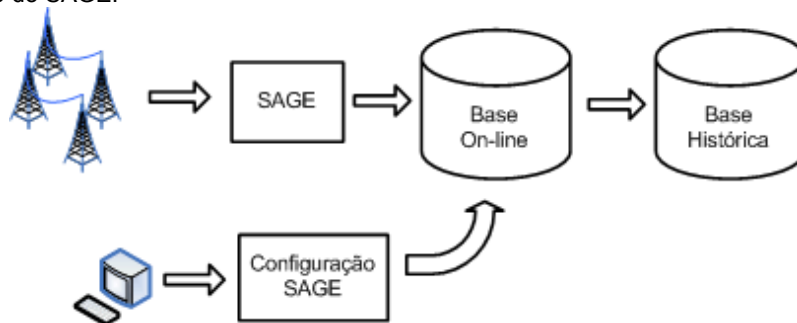


Figura 2: Fluxo da informação no SAGE

A base histórica do SAGE é responsável pelo armazenamento histórico da operação do sistema elétrico. Armazena tanto informações provenientes da operação do sistema elétrico, quanto provenientes da evolução deste. Suas estruturas basicamente armazenam séries temporais e são gerenciadas pelo GBH, sistema de Gerenciamento da Base Histórica. O GBH é configurável e permite que qualquer atributo da base on-line seja historiado.

2.2 Processo de carga a base histórica

O procedimento de carga da base histórica é feito através de uma aplicação, STI_CARGBH, que se conecta simultaneamente à base on-line do SAGE e a um banco de dados relacional, e que grava no banco relacional os dados provenientes do tempo real conforme mostra a figura 3.

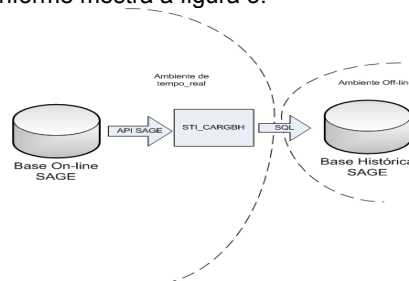


Figura 3 Processo de carga da base histórica.

O aplicativo possui um sofisticado sistema de salvaguarda que garante o armazenamento dos dados históricos a despeito de falhas na conexão com o banco de dados e também executa a função de arquivamento, permitindo a construção de um “arquivo morto” com dados históricos da operação do sistema elétrico, em formato portátil e que pode ser armazenado em mídias específicas para esse fim.

2.3 Estrutura da base histórica do SAGE

Na base histórica, as entidades são organizadas em dois tipos de tabela: **referência** e **dinâmicas**. A figura 4 mostra o relacionamento entre uma entidade da base on-line e suas respectivas tabelas de referência e dinâmicas.

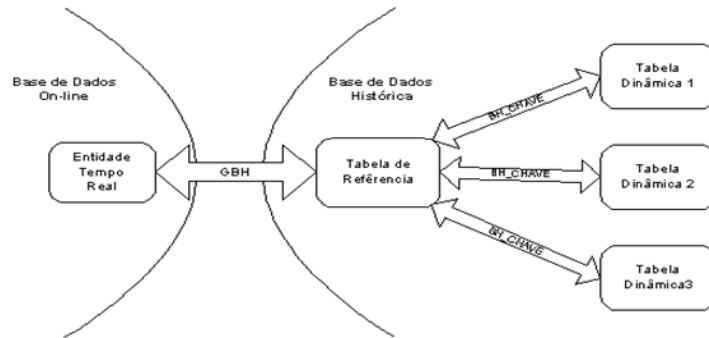


Figura 4 - Relação entre as entidades da base on-line e as tabelas da base histórica

As tabelas de referência têm como função manter um histórico da configuração do sistema elétrico. Ela armazena as informações cadastrais e sua evolução. Seus atributos representam uma determinada característica do sistema tal como unidade de medida, tipo de equipamento, identificador, etc. As entidades da base on-line só podem ter uma e somente uma tabela de referência na base histórica. Tabelas dinâmicas armazenam as informações provenientes do processamento do sistema como, por exemplo, valores medidos, estados digitais, etc. Seus atributos representam valores do sistema que variam em tempo real. Uma importante característica das tabelas dinâmicas é que elas crescem continuamente, o que leva à necessidade, por questão de desempenho, de estabelecer políticas de particionamento em várias tabelas físicas cada uma representando um período de tempo (dia, semana, mês). Outra característica importante é a forma com que são gravados os dados nas tabelas dinâmicas que denominamos esquema de gravação. Os esquemas de gravação podem ser classificados em dois grupos: síncrono e assíncrono. No esquema síncrono os valores do tempo real são varridos periodicamente e no esquema assíncrono a varredura se dá sempre que a base tempo-real identifica que ocorreram mudanças nos dados relacionados.

2.4 Processo de configuração da Base Histórica

Como foi dito anteriormente a estrutura da base histórica é configurável de forma a permitir que qualquer informação da base on-line do SAGE possa ser historiada. A configuração é feita através da seleção das entidades/atributos da base on-line que serão historiados. A partir dessa configuração são geradas as estruturas da base histórica e o programa STI_CARGBH, responsável por preenchê-las. As Informações sobre a estrutura da base histórica são armazenadas na própria base através de tabelas específicas, formando um conjunto de metadados que serão muito úteis no processo de evolução da BDH para um *data warehouse*. A figura 5 ilustra o processo de configuração da base histórica do SAGE.

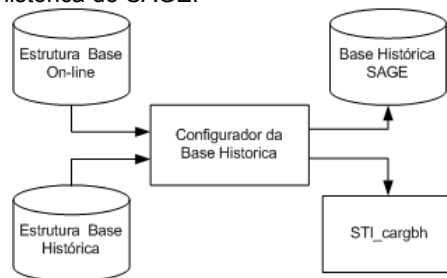


Figura 5: Processo de configuração da base histórica do SAGE

2.5 Observações

Apesar de muito eficiente no processo de captura e armazenamento dos dados de tempo real, a solução fornece poucas facilidades para a exploração dos dados por parte dos usuários finais. Algumas iniciativas foram tomadas nesse sentido, tal como o processo de consolidação que além de excluir medidas inválidas executa algum pós-processamento através do cálculo de valores máximos, médios e mínimos de medidas analógicas em diferentes precisões (horária, diária, etc.). Outro aspecto importante é o caso da empresa que possui mais de um historiador, seja porque sua operação está dividida entre vários centros de controle ou se por estratégia de implantação o histórico é feito nos níveis hierárquicos inferiores da estrutura de supervisão e controle. Nesse caso o sistema não apresenta nenhuma solução para integração desses históricos. A aplicação de técnicas de *data warehousing* vem como resposta a essas questões como veremos a seguir.

3.0 - DATA WAREHOUSE

Um *data warehouse* é um sistema computacional utilizado para armazenar informações relativas às atividades de uma organização em bancos de dados, de forma consolidada. O desenho da base de dados é orientado ao usuário final, privilegiando o acesso à informação. Sua utilidade é servir de fonte de consulta para aplicativos voltados à análise de grandes volumes de dados e à obtenção de informações estratégicas que podem facilitar a tomada de decisão. O sistema pode ser decomposto em quatro partes, cada qual com uma responsabilidade específica. As quatro partes que compõem um sistema de *data warehouse* são:

- Subsistema de captura de dados - fonte de informações do sistema.
- Área de armazenamento centralizado de dados brutos (*data staging area*).
- Área de apresentação de informações consolidadas.
- Ferramentas avançadas para análise e mineração de dados (*data-mining & analitic tools*).

O subsistema de captura de dados é responsável pela interface com o processo, no nosso caso o processo de operação do sistema elétrico. A base histórica do SAGE se encaixa nesse contexto como uma ferramenta que permite a gravação em um banco de dados relacional das séries temporais representativas da operação/evolução do sistema elétrico. A segunda parte de um *data warehouse* é composta por um repositório intermediário, conhecido na literatura como *Data Staging Área*. Nesse repositório são armazenados dados das diversas fontes de informações do subsistema de captura onde são consolidados, manipulados e combinados e finalmente entregues à área de apresentação do sistema, terceira parte do sistema. Esse procedimento é conhecido como ETL, do inglês *Extract Transform Load*. A área de apresentação é onde a informação é disponibilizada para o usuário final. Os dados são organizados através de uma modelagem dimensional como veremos mais adiante. A última parte é composta por aplicativos para análise e mineração das informações disponíveis na área de apresentação. Existem várias ferramentas e algoritmos disponíveis para atender essa parte do sistema, a qual deve ser modelada a partir das necessidades do usuário final e está fora do escopo deste trabalho. A figura 6 ilustra a estrutura de um *data warehouse* de informações do sistema elétrico baseado na base histórica do SAGE. Note que a arquitetura prevê a integração de várias fontes de dados o que permite a geração de um *data warehouse* a partir de várias bases históricas dos sistemas SAGE.

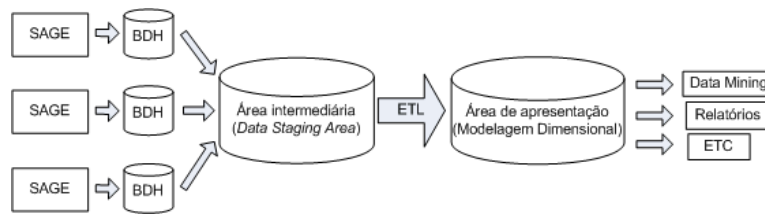


Figura 6: Estrutura de um data warehouse

3.1 Modelagem Dimensional

Modelagem dimensional é uma técnica para projeto de bancos de dados que busca apresentar os dados de uma forma padronizada através de uma estruturação intuitiva e que permite acessos a grandes volumes de dados com alto desempenho. O modelo dimensional é composto basicamente de duas classes de objetos, denominadas tabelas de fatos e tabelas de dimensões. A construção básica é formada por uma tabela de fatos que é composta de valores medidos associados através de chaves estrangeiras às tabelas de dimensão, como ilustra a figura 7. A construção é denominada *star-join* devido à forma de estrela característica do modelo.

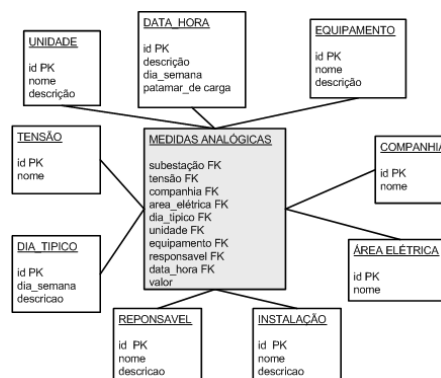


Figura 7: Construção básica star model

A tabela de fatos é composta por várias chaves estrangeiras que expressam relacionamentos de cardinalidade M:N entre as tabelas de dimensão. Uma tabela de fatos contém uma ou mais medições, ou fatos, que ocorrem para uma determinada combinação de chaves que define cada registro. Na figura 7 a tabela de fatos

MEDIDAS_ANALOGICAS, o fato é o valor medido. Os “fatos” em uma tabela de fatos devem ser numéricos e agregáveis. As funções de agregação são a alma do modelo, pois uma aplicação típica de *data warehouse* raramente busca um único registro por vez, normalmente as consultas retornam milhares, até milhões de registros que são selecionados e agrupados através de critérios definidos pelas chaves de ligação com as tabelas de dimensão e seus fatos são agregados através de funções do tipo somatório, média, etc. As tabelas de dimensão por sua vez contêm uma chave primária e informações descritivas normalmente na forma de textos. Seus atributos são utilizados como filtros primários para seleção dos fatos que comporão o resultado de uma consulta. Por exemplo, na estrutura da figura 7 a consulta da geração total por companhia em um determinado período pode ser feita de forma simples e intuitiva. Supondo que a chave 15 obtida através de uma na tabela dimensional UNIDADE indica valores de geração e os valores 5789 e 6397 referem-se às datas de início e fim do período desejado obtidas na dimensão DATA_HORA. A query teria a seguinte forma:

```
SELECT COMPANHIA, SUM (VALOR)
FROM MEDIDAS_ANALOGICAS
WHERE UNIDADE = 15 AND DATA_HORA >= 5789 AND DATA_HORA <= 6397
GROUP BY COMPANHIA;
```

O exemplo deixa claro que a qualidade de uma modelagem está diretamente atrelada a qualidade e abrangência das tabelas de dimensão. A melhor forma de se consultar as tabelas de fatos é através das chaves estrangeiras que devem ser obtidas previamente com consultas às tabelas de dimensão. Fazer consultas que envolvam joins entre as tabelas de fatos e dimensão, apesar de não serem proibidas terão um impacto negativo no tempo de resposta. Outro aspecto importante é a necessidade do desenvolvimento de funções de agregação adequadas às necessidades das aplicações voltadas ao setor elétrico. Um *data warehouse* típico é composto por várias estruturas semelhantes à da figura 7 como é mostrado na figura 8. Cada estrela é composta uma tabela de fatos, cercada por várias tabelas de dimensão. Uma tabela de dimensão pode ser compartilhada entre várias estrelas premissa fundamental para operações do tipo *drilling-across* que veremos a seguir.

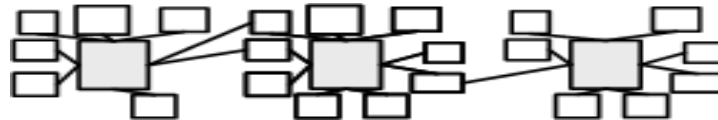


Figura 8: Estrutura típica de um *data warehouse* “star model”

3.2 Operações básicas em um *data warehouse*

Apesar do *data warehouse* ser um conceito genérico aplicável em diversas áreas, todos devem conter alguns conceitos fundamentais comuns. Os três conceitos mais importantes referem-se às operações de *drilling-down*, *drilling-across* e o tratamento temporal das informações. Aplicações modernas de *data warehousing* incorporaram esses conceitos tão profundamente que estes virtualmente tornaram-se condição necessária para que um sistema de armazenamento de dados seja considerado realmente um *data warehouse*.

3.2.1 Drilling-down

Drilling down, que significa perfurar em inglês, é a operação mais fundamental em um *data warehouse*. Através dela podemos analisar as informações agregadas em vários níveis de consulta apenas adicionando ou removendo atributos na query. Por exemplo, considerando a estrutura da figura 7 conseguimos obter a geração por companhia através da query:

```
SELECT COMPANHIA, SUM (VALOR)
FROM MEDIDAS_ANALOGICAS
WHERE UNIDADE = 15 AND DATA_HORA >= 5789 AND DATA_HORA <= 6397
GROUP BY COMPANHIA;
```

Caso se deseje a geração total por cada área elétrica monitorada pelas companhias, basta adicionar a chave estrangeira referente à dimensão AREA_ELÉTRICA à consulta obtendo a seguinte query:

```
SELECT COMPANHIA, AREA_ELÉTRICA, SUM (VALOR)
FROM MEDIDAS_ANALOGICAS
WHERE UNIDADE = 15 AND DATA_HORA >= 5789 AND DATA_HORA <= 6397
GROUP BY COMPANHIA, AREA_ELÉTRICA;
```

Da mesma forma podemos fazer a consulta por instalação bastando adicionar a chave para a dimensão INSTALACAO à consulta. A capacidade de especialização/generalização de forma simples e intuitiva é a essência do conceito de *drilling-down*.

3.2.2 Drilling-across

Através desse conceito é possível agregar medidas de duas ou mais tabelas de fatos distintas através das dimensões comuns. É importante salientar a importância da correspondência semântica entre as medidas a serem

agregadas nas diversas tabelas de fatos que compõe a operação, sobre a pena de estarmos somando “bananas” com “laranjas”.

3.2.3 Handling-time

Todos os dados em um *data warehouse* têm que ter claramente definido um período de validade. Se a estrutura de uma entidade pode mudar no decorrer do tempo é necessário que seja possível associar cada versão da entidade com as versões das demais entidades contemporâneas. Em outras palavras um *data warehouse* deve permitir a recuperação de todo o contexto do objeto de negócio em um determinado período do tempo. No caso do sistema elétrico é válido dizer que os valores medidos em um determinado período estão intrinsecamente ligados à topologia do sistema elétrico no período em estudo. Isso é feito através da inclusão de atributos de início e fim de operação de um determinado recurso do sistema na respectiva dimensão. O tratamento temporal das tabelas de dimensão é um dos pontos mais críticos no desenvolvimento de um *data warehouse* principalmente no que tange ao tratamento de alterações de informações em tabelas dimensionais, que se caracterizam por um crescimento lento quando comparado ao crescimento das tabelas de fatos.

Sob o ponto de vista das tabelas de fatos, um aspecto importante a se ressaltar é que um *data warehouse* deve suportar as formas naturais de se observar um processo ao longo do tempo, ou seja, através de eventos ocasionais, varreduras periódicas e a obtenção do estado corrente. A modelagem dimensional proporciona um ambiente muito conveniente para atender a essa questão uma vez que sua estrutura central, a tabela de fatos, é composta de medidas que são feitas em um determinado instante do tempo. Em um *data warehouse* clássico a data de um determinado evento é representada como mais uma dimensão associada ao fato através de uma chave estrangeira. A literatura sugere a criação de tabelas dimensionais contendo um calendário que é expandido por um determinado período, que em sua estrutura teriam atributos com informações que seriam muito úteis durante a análise de dias típicos e outros eventos periódicos. Para sistemas com uma periodicidade diária essa técnica é perfeitamente viável, pois para armazenar dados para um período de, por exemplo, cinquenta anos, seria necessária uma tabela dimensional de $50 \times 365 = 18250$ registros o que é perfeitamente factível, para uma precisão horária seriam necessários 438000 registros o que já é uma relação de um tamanho considerável, porém, ainda factível. No caso de um *data warehouse* com dados provenientes da operação elétrica a precisão é na casa de segundos e em alguns casos milissegundos o que inviabiliza a amarração da data/hora do evento através de uma tabela dimensional. A solução é a utilização de campos do tipo data/hora nativos do banco de dados na própria tabela de fatos. Essa solução não impede a criação de uma tabela dimensional calendário com precisão horária para associarmos aos fatos, ela seria muito útil na análise do comportamento do sistema elétrico em, por exemplo, períodos típicos ou patamares de carga.

3.3 Cubos OLAP

Outra estrutura utilizada pela tecnologia são os cubos multidimensionais OLAP. Inicialmente indicados para pequenos volumes de dados os cubos OLAP permitem que ferramentas de análise efetuem cálculos e comparações bastante sofisticados de forma simples e intuitiva. A literatura recomenda a construção dos cubos a partir dos dados já organizados em um esquema estrela através de ferramentas. A combinação das duas estruturas forma um ambiente bastante poderoso para integração de ferramentas de análise e mineração de dados. No escopo desse trabalho não será aprofundado o uso de tais cubos, pois, consideramos que este seria um segundo passo na evolução dos bancos de dados históricos da operação a ser apresentado em trabalho futuro.

3.4 Observações

Como podemos perceber a tecnologia de *data warehousing* com sua a modelagem dimensional orientada à construção de relacionamentos do tipo estrela (*star joins*), é um caminho viável para evolução de bancos de dados históricos da operação elétrica no que se refere à tarefa de disponibilizar informações consolidadas ao usuário final, cabendo aqui algumas observações:

- Sua estrutura de forma padronizada fornece um ambiente amigável para o desenvolvimento de procedimentos ETL automatizados para a importação de dados.
- A versatilidade de sua estruturação permite o aprimoramento contínuo sem que o contexto histórico seja perdido.
- Serão necessárias algumas adaptações no modelo para atender os requisitos de precisão, especialmente no que se refere ao tratamento de datas.
- O projeto seria muito enriquecido com o desenvolvimento de funções de agregação especializadas para o tratamento de medidas elétricas.
- O sucesso da empreitada passa necessariamente pelo bom desempenho da operação ETL para a atualização das estruturas.
- Existem no mercado ferramentas gráficas para geração automatizada de procedimentos ETL.

No capítulo seguinte será mostrada uma proposta de *data warehouse* gerada a partir da base de dados histórica do SAGE.

4.0 - A BASE HISTÓRICA DO SAGE E *DATA WAREHOUSE*

Comparando a estrutura da base histórica do SAGE baseada em tabelas dinâmicas e de referência, com a estruturação proposta pela modelagem dimensional verificamos de imediato uma correlação direta entre as tabelas de referência e as dimensões e também entre as tabelas dinâmicas e as tabelas de fatos. A descrição das estruturas das bases de dados do SAGE, disponível para consulta no próprio banco de dados, contribui para a geração automática dos procedimentos ETL e para atualização da área de apresentação a partir dos dados provenientes da base histórica do SAGE.

4.1 Tabelas de Referência X Dimensões

Apesar de existir uma correlação direta entre as tabelas de referência e as dimensões elas estão estruturadas de forma distinta. As tabelas de referência da base histórica do SAGE são organizadas no modelo entidade-relacionamento, e as tabelas de dimensão no modelo dimensional, o que nos obrigará a adotar algumas regras de transformação. Por exemplo:

- Toda tabela de referência associada diretamente a uma tabela dinâmica deve gerar pelo menos uma dimensão contendo sua chave primária e atributos sem domínio conhecido.
- Os atributos relativos ao período de operação do recurso (datas de início e fim, BH_INDTR e BH_SINONIMO) devem ser replicados na tabela de dimensão. As tabelas de referência já utilizam o método dois para tratamento de alterações em tabelas de dimensão descrito em [14].
- Para toda chave estrangeira associada a outra tabela de referência gerar uma dimensão com os dados da tabela referenciada. Essa operação deve ser efetuada recursivamente repetindo todo o tratamento para todas as tabelas de referência envolvidas.
- Os atributos com domínio definido devem gerar uma nova dimensão contendo todos os valores possíveis.

Além das dimensões criadas pelas regras acima, devem ser criada a uma dimensão para calendário com precisão horária, uma de dias típicos e no caso do *data warehouse* ser gerado a partir de várias fontes de dados uma dimensão para indicar a responsabilidade pela informação.

4.2 Tabelas dinâmicas X Fatos

A tabela de fatos é composta de todos os atributos das tabelas dinâmicas acrescidos de chaves estrangeiras apontando para todas as dimensões geradas a partir da tabela de referência associada, e para as dimensões de calendário, dias típicos e responsabilidade (para o caso de mais de um SAGE alimentando a *data warehouse*).

Um aspecto importante é a possibilidade de derivarmos várias tabelas dimensionais a partir de uma tabela dinâmica. As tabelas de fatos geradas variam em precisão e o horizonte de tempo armazenado em cada uma. Por exemplo, a partir da tabela dinâmica de medidas analógicas podemos gerar dimensões com precisão de segundos, horária, diária, mensal e anual, cada qual com um horizonte que melhor se adapte às necessidades do usuário final. As tabelas de fatos podem conter valores médios, máximos e mínimos no período.

4.3 Procedimentos ETL

Os procedimentos ETL devem ser executados periodicamente, inicialmente com uma frequência horária. Esses procedimentos são responsáveis por manter atualizadas as estruturas que compõe o *data warehouse*. Para isso o ETL deve executar as seguintes tarefas:

- Verificar mudanças na estruturas da base histórica do SAGE e aplicá-las no *data warehouse*.
- Replicar as atualizações das tabelas de referência nas tabelas de dimensão.
- Carregar nas tabelas de fatos os dados das tabelas dinâmicas.

Tendo em vista que a estrutura da base histórica do SAGE é descrita em seu catálogo, é possível que com algumas adaptações os procedimentos de ETL sejam gerados a partir dessa definição de forma automática.

4.4 Retratos do tempo real

Uma funcionalidade muito importante é a vista do processo elétrico em tempo real. Ela permite consultas do estado corrente do sistema, constituindo-se em uma importante interface para aplicações externas ao centro de controle, tal como portais WEB. A estratégia para o desenvolvimento dessa funcionalidade é criar para cada tabela dinâmica uma tabela de fatos que só armazena o valor corrente das medidas. A taxa de atualização é um fator importante, o que leva à necessidade de procedimentos ETL específicos e independentes dos procedimentos ETL de carga de dados históricos.

4.5 Observações

Apesar da correlação encontrada entre as tabelas de referência e dinâmicas da base histórica do SAGE e as tabelas de fatos e dimensão do modelo dimensional, a conversão entre os modelos não é trivial. Existe a necessidade de aprimoramentos nos catálogos das bases histórica e on-line do SAGE, entre os quais a necessidade de se descrever os relacionamentos entre as tabelas de referência para o procedimento de geração das tabelas de dimensionais, além de melhorias e padronização nas definições dos domínios de alguns atributos.

5.0 - CONCLUSÃO

A tecnologia de *data warehousing* com sua modelagem dimensional constitui-se em um ambiente extremamente poderoso para o armazenamento e difusão de informações históricas provenientes da operação e expansão do sistema elétrico. O casamento dessa tecnologia com o SAGE, através de sua base histórica é uma importante agregação de valor ao produto e que impactará de forma positiva e significativa o uso dos dados históricos da operação, possibilitando o desenvolvimento de aplicações mais eficientes orientadas ao usuário final, auxiliando no apoio à decisão, análise de contingências pós-operação e planejamento da expansão do sistema elétrico determinando de forma mais precisa as necessidades de investimento.

A capacidade de agregação em um repositório único dos dados históricos produzidos pelos diversos SAGES instalados na empresa permite uma visão global do sistema elétrico e se mostra uma arquitetura preparada para acompanhar a evolução do sistema elétrico sem a necessidade de investimento significativo em infra-estrutura de armazenamento de dados tais com licenças de bancos de dados, servidores, etc.

Uma vez implantado, o ambiente para desenvolvimento proporcionado pelo data warehouse servirá de forma muito mais efetiva como fonte de informações para todos os setores envolvidos com a atividade fim de uma empresa de energia elétrica, aumentando de forma global a qualidade final de seus processos e proporcionando serviços de melhor qualidade em uma área extremamente estratégica para o crescimento do país.

6.0 - REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ZARUR, PEDRO DANIEL - Supervisão e Controle em Tempo Real de Sistemas Elétricos baseada em Computação Grid [Rio de Janeiro] 2005 IX, 128 p. 29,7 cm (COPPE/UFRJ, M.Sc., Engenharia Elétrica, 2005).
- [2] Data Warehouse Middleware Towards an Unified Resource Management, P. Assis - DEE, ISEP, Instituto Politécnico do Porto, 4200-485 Porto, Portugal e J. A. Martins - DET, Universidade de Aveiro, 3810-193 Aveiro, Portugal.
- [3] Experimental Evaluation of a New Distributed Partitioning Technique for Data Warehouses - Jorge Bernardino - Institute Polytechnic of Coimbra, Henrique Madeira - University of Coimbra DEI, Portugal.
- [4] Research on the SCADA /EMS System Data Warehouse Technology - Xiaofeng He, Gang Wang, Jiancang Zhao - 2005 IEEE/PES Transmission and Distribution Conference & Exhibition: Asia and Pacific Dalian, China
- [5] A Dimensional Modeling Manifesto – Ralph Kimball - Agosto 1997
- [6] Manual da Base Histórica do SAGE – Pedro Daniel Zarur – Novembro 2006
- [7] Data Ware House Fundamentals: TCO Starts With the End User – Ralph Kimball – Maio 2003
- [8] Data Ware House Fundamentals: An Engineer's View – Ralph Kimball – Julho 2003
- [9] Data Ware House Fundamentals: Design Constraints and Unavoidable Realities – Ralph Kimball – Setembro 2002
- [10] Data Ware House Fundamentals: Two Powerful Ideas Realities – Ralph Kimball – Setembro 2002
- [11] Data Ware House Fundamentals: Fact Tables and Dimension Tables – Ralph Kimball – Janeiro 2003
- [12] Data Ware House Fundamentals: The Soul of the Data Warehouse, Part One: Drilling Down – Ralph Kimball – Março 2003
- [13] Data Ware House Fundamentals: The Soul of the Data Warehouse, Part Two: Drilling Across – Ralph Kimball – Abril 2003
- [14] Data Ware House Fundamentals: The Soul of the Data Warehouse, Part 3: Handling Time – Ralph Kimball – Abril 2003
- [15] Data Ware House Designer: Realtime Partitions - Ralph Kimball – Fevereiro, 2002

7.0 - DADOS BIOGRÁFICOS

Pedro Daniel Zarur

Nascido no Rio de Janeiro, RJ em 12 de abril de 1962.

Mestrado (2005) COPPE-UFRJ e Graduação (1987) em Engenharia Elétrica: Universidade Gama Filho

Empresa: CEPEL – Centro de Pesquisas de Energia Elétrica, desde 2006.

Pesquisador do Departamento de Automação de Sistemas

Luiz Antonio Cordeiro Pereira

Nascido no Rio de Janeiro, RJ em 09 de julho de 1959.

Mestrado (1991) em Sistemas de Computação - COPPE-UFRJ, Graduação (1981) em Engenharia Elétrica: UFRJ

Empresa: CEPEL – Centro de Pesquisas de Energia Elétrica, desde 1985.

Pesquisador do Departamento de Sistemas Elétricos