

# Ferramenta de Auxílio no Controle de Perdas Reais de Água

Ricardo Martins Carvalho  
Instituto Federal da Bahia  
Salvador, Brasil  
Email: ricardo.carvalho@ifba.edu.br

Jowaner de Oliveira Araújo  
Instituto Federal da Bahia  
Salvador, Brasil  
Email: jowaner@ifba.edu.br

**Resumo**—Este trabalho apresenta a ferramenta de auxílio no controle de perda reais de água. Inicialmente serão abordados os principais conceitos e definições sobre perdas, englobando as perdas comerciais e perdas reais de água decorrentes de vazamentos. Visando a dificuldade que existe em se controlar a quantidade de água perdida durante ocorrências de vazamentos, este estudo irá levar em consideração o tempo de ocorrência de vazamento desde a solicitação de um serviço de reparo por, clientes ou internamente, até a execução da sua correção. A análise dessa problemática foi feita especificamente, baseando-se em reclamações feitas por clientes situados na Localidade de Pirajá, pertencente à Unidade Regional de Pirajá (UMJ). Esta unidade fica localizada na Região Metropolitana de Salvador (RMS) que pertence à Diretoria Metropolitana (DM) da Embasa. Para tanto, foi desenvolvido um *data warehouse* disponibilizando informações sobre o desperdício de água, utilizando recursos gráficos, visando auxiliar o processo de tomada de decisões.

**Keywords**—Perdas, vazamentos, vazão, zona de abastecimento, unidade regional, *data warehouse*, Embasa.

## I. INTRODUÇÃO

A água é um elemento essencial para vida, porém, está se tornando um bem cada vez mais escasso. Cerca de 97% de toda a água existente em nosso planeta é salgada e considerada imprópria para consumo humano. Da parcela de água restante, apenas 0,03% está diretamente disponível para consumo humano [1]. O aumento da demanda de consumo de água e a constante degradação dos mananciais também contribuem para esta escassez. A redução do consumo e a preservação dos mananciais, portanto, são medidas essenciais para combater a escassez da água e o colapso dos sistemas de abastecimento [1].

Além das questões supracitadas, um terceiro problema afeta a disponibilidade dos recursos hídricos, o volume de perdas hídricas identificado nos Sistemas de Abastecimento de Água<sup>1</sup> (SAA's) brasileiros. Do ponto de vista físico, as perdas são causadas tanto por falhas nos sistemas (vazamentos, transbordamentos, etc.), quanto pela ocorrência de consumo não registrado.

Ademais, imprecisões nos métodos de mensuração, decorrentes de falhas dos hidrômetros ou estimativas pouco fidedignas, comprometem o cálculo das perdas (comerciais) das concessionárias de abastecimento. Com índices altos de

perda, os SAA's passam a demandar mais água do que o necessário para consumo final, tendendo assim a agravar o problema da escassez de água.

De acordo com Tsutiya (2004) [2], perda hídrica é toda água tratada que foi produzida e se perdeu no caminho, não chegando ao uso final pelos clientes da companhia de saneamento.

Para os SAA's, conceituam-se como perdas hídricas a diferença entre o volume de água disponibilizado após o tratamento e o volume total registrado como consumo, para um dado período de tempo. A apuração dos índices de perdas é tradicionalmente realizada por meio do cálculo da diferença entre macromedição<sup>2</sup> e micromedição<sup>3</sup>, sendo estes números reflexo da parcela do volume de água que não foi contabilizado como consumo final [2][3][4].

Este trabalho objetiva analisar o balanço hídrico numa empresa de saneamento, com foco nas perdas reais de água, propondo identificar a quantidade de água perdida durante ocorrências de vazamentos. Logo, será avaliado o tempo entre a primeira notificação do cliente até seu atendimento final, visto que, com informações confiáveis, a empresa tende a adotar ações assertivas de mitigação ou eliminação dessas perdas. Como objetivos específicos, teremos:

- 1) descrever conceitos e definições sobre balanço hídrico, no contexto de saneamento, e perdas físicas e comerciais de água;
- 2) Elaboração do modelo para analisar as reclamações de vazamentos, em ramais, de clientes de uma localidade, pertencente ao município de Salvador - BA;
- 3) Quantificar o desperdício de água oriundo desses vazamentos através dos resultados obtidos a partir da análise;
- 4) Elaboração de gráficos e mapas para visualização dos dados analisados.

Uma parte significativa das redes de distribuição de água são subterrâneas, tornando complexa a localização de vazamentos, e, em certos casos, necessitam utilizar meios destrutivos durante a sua descoberta (Figura 1). Consequentemente, por serem bastante rotineiras, as inspeções acabam por encarecer o monitoramento e localização dos vazamentos.

<sup>1</sup>Estruturas complexas, responsáveis por captar a água dos mananciais e distribuí-la diretamente aos reservatórios dos consumidores em condições adequadas para o consumo humano.

<sup>2</sup>Medição feita na saída dos grandes reservatórios distribuídos pela cidade, para determinar a quantidade de água tratada disponibilizada.

<sup>3</sup>Medição da quantidade de água consumida contabilizada pelos hidrômetros, nas ligações dos clientes das empresas de saneamento.



Figura 1. Vazamento não aparente.

No setor de saneamento básico, é bastante evidente a necessidade de se explorar e aplicar métodos de rastreamento de vazamentos, principalmente por meio de técnicas eficazes e menos onerosas do que as inspeções físicas tradicionais. Tecnologias para rastreamento de vazamentos ainda são ineficazes. Os mais indicados para tal objetivo são os equipamentos baseados em propagação de ondas sonoras, tais como o geofones<sup>4</sup>, utilizados por algumas empresas, visando auxiliar no rastreamento deste tipo de problema. Todavia, é opinião frequente dos técnicos sanitaristas que tais equipamentos, além de caros e de difícil operação, são restritivos quanto às condições de funcionamento e de instalação para garantirem precisão nos resultados [5].

Além desta Introdução, este trabalho está dividido nas seguintes seções:

- a Seção II apresenta assuntos que nortearão a progressão deste trabalho;
- A Seção III engloba o processo que será executado para construção da solução proposta;
- A Seção IV compreende o contexto atual da empresa utilizada para o estudo deste trabalho, a aplicação da metodologia proposta e os resultados obtidos com uso solução proposta;
- A Seção V conclui este trabalho, destacando as restrições e perspectivas futuras;
- A Seção VI contém apêndices, com funcionalidades utilizadas para este trabalho;
- A Seção VII contém o anexo da solicitação e da autorização para uso dos dados da Embasa;
- Por fim, a Seção VIII traz os agradecimentos.

## II. REFERENCIAL TEÓRICO

Esta seção apresenta os principais assuntos relacionados a este trabalho. A subseção II-A relata sobre a Empresa Baiana

<sup>4</sup>Instrumento que serve para escutar ruídos na terra, utilizando-se da prospecção através do eco.

de Águas e Saneamento S.A (Embasa). A subseção II-B define o balanço hídrico. A subseção II-C, discorre sobre Sistemas de Abastecimento. Na subseção II-D, é abordado o assunto de Perdas. Na subseção II-E será apresentado as ferramentas e tecnologias utilizadas neste trabalho.

### A. Empresa Baiana de Águas e Saneamento S.A. - Embasa

A Embasa é a empresa responsável pelo tratamento e distribuição de água, assim como coleta, tratamento e destinação do esgotamento sanitário, dos municípios onde atua, no estado da Bahia [6]. Hoje a Embasa atua em 366 municípios, conforme demonstra Figura 2, disponibilizada pela Embasa [6], nesta figura os grupos coloridos representam as Unidades Regionais da Embasa enquanto os grupos em branco representam os municípios que possuem abastecimento de água independente:

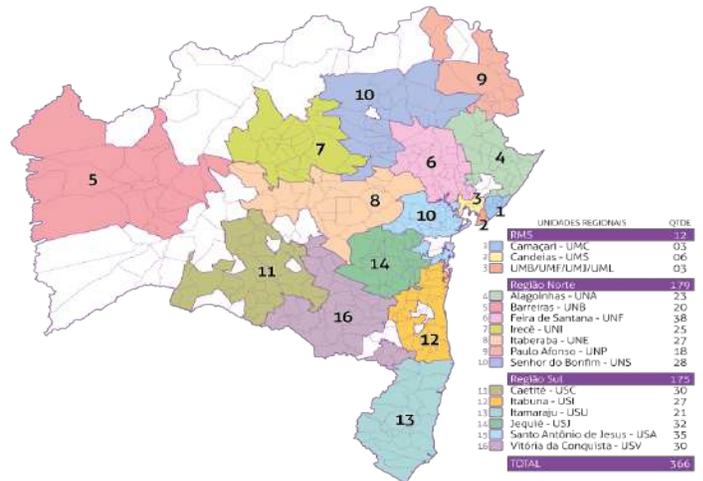


Figura 2. Municípios abastecidos pela Embasa.

De acordo com informações disponibilizadas no site da Embasa, a empresa é dividida em sete diretorias (Figura 3) [6], sendo que as Diretorias de Operação da RMS e do Interior são responsáveis pelas operações que envolvem, desde o tratamento até a distribuição de água para os clientes. Estas diretorias subdividem-se em Superintendências responsáveis pela produção e serviços de água e esgoto entre as quais estão a Superintendência Metropolitana (MS), do Norte (IN) e do Sul (IS). Abaixo das Superintendências encontram-se as Unidades Regionais, que correspondem à distribuição com base nos polos Administrativos. São ao todo 19 Unidades Regionais<sup>5</sup> conforme descritos a seguir:

- 1) RMS: UMB, UMF, UMJ, UML (as quatro primeiras dentro do município de Salvador), UMC e UMS;
- 2) NORTE: UNA, UNB, UNE, UNF, UNI, UNP e UNS;
- 3) SUL: USA, USC, USI, USJ, USU e USV.

As Unidades Regionais subdividem-se em localidades, que correspondem a municípios inteiros ou divisões de municípios (quando estes são muito grandes), povoados e distritos. A Figura 4, cedida pela Embasa [6], apresenta as 6 localidades

<sup>5</sup>As duas letras iniciais, utilizadas como siglas das Unidades Regionais, significam: U - Unidade de Negócio ou Unidade Regional e a segunda letra (M,N,S) corresponde a região (Metropolitana, Norte e Sul). A terceira letra corresponde ao nome da unidade.

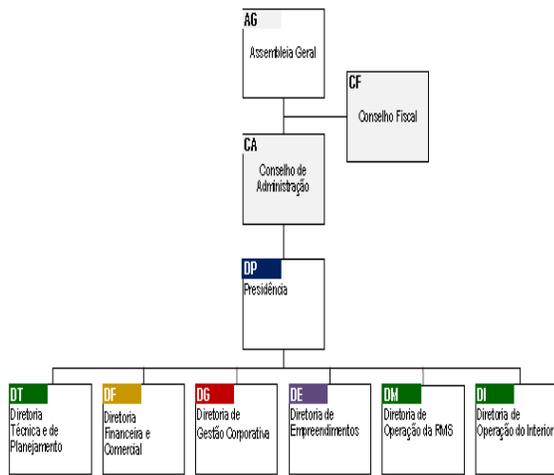


Figura 3. Organograma - Diretorias Embasa.

que compõe o abastecimento de Salvador, Lauro de Freitas e Simões Filho, sendo formada pelas localidades:

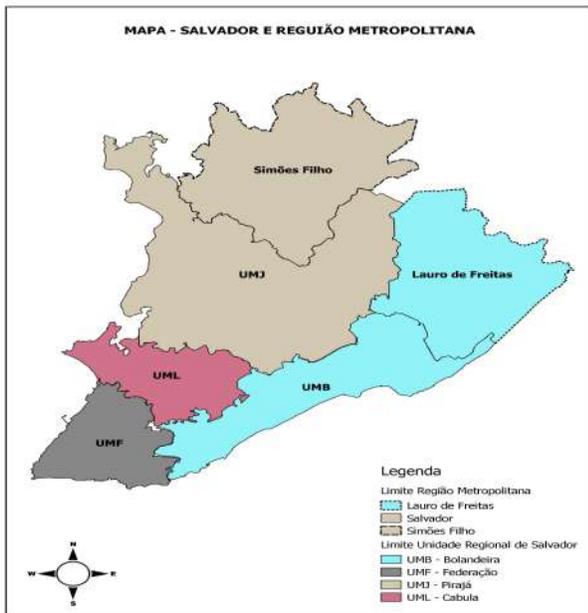


Figura 4. Unidades Regionais pertencentes a MS.

- 1) Bolandeira (UMB);
- 2) Federação (UMF);
- 3) Cabula (UML);
- 4) Pirajá (UMJ);
- 5) Lauro de Freitas (UMB);
- 6) Simões Filho (UMJ).

Com o intuito de melhorar o processo de distribuição de abastecimento, as localidades ainda possuem divisões chamadas de setores de abastecimento. Esta separação permite um melhor controle e execução de ações específicas [7].

### B. Balanço Hídrico

O termo 'Balanço Hídrico' refere-se, primeiramente, ao controle de água, estudo de climas, zoneamento de áreas de

maior incidência de chuvas, regiões com maior deficiência desse recurso, quantidade de água que está disponível em solo, assim como controle da desidratação no corpo humano (na área de Saúde), entre outros [8].

Ao que tange serviços de abastecimento de água, este termo se refere ao controle de água tratada disponibilizada a população e problemas que afetam os serviços de abastecimento. No setor de saneamento, o combate aos vazamentos é um dos problemas mais comuns.

### C. Sistema de Abastecimento de Água (SAA)

Sistema de abastecimento de água consiste num conjunto de elementos, mecanismos e serviços, com o propósito de captar água de mananciais e rios, executar o tratamento desta água, quando necessário e distribuir a água para os consumidores [5]. A Figura 5, disponível no site da Copasa<sup>6</sup> [9], apresenta o funcionamento de um SAA. Este processo envolve captação, tratamento e distribuição. Conforme esta figura, o estudo deste trabalho considera o ponto de partida os reservatórios de água tratada, onde há a primeira medição (macromedição) da água antes da distribuição à população. Esta água percorre pelas adutoras<sup>7</sup>, que distribuem águas entre outros reservatórios e redes de distribuição de água<sup>8</sup> que utilizam os ramais<sup>9</sup> para abastecimento de água dos clientes.

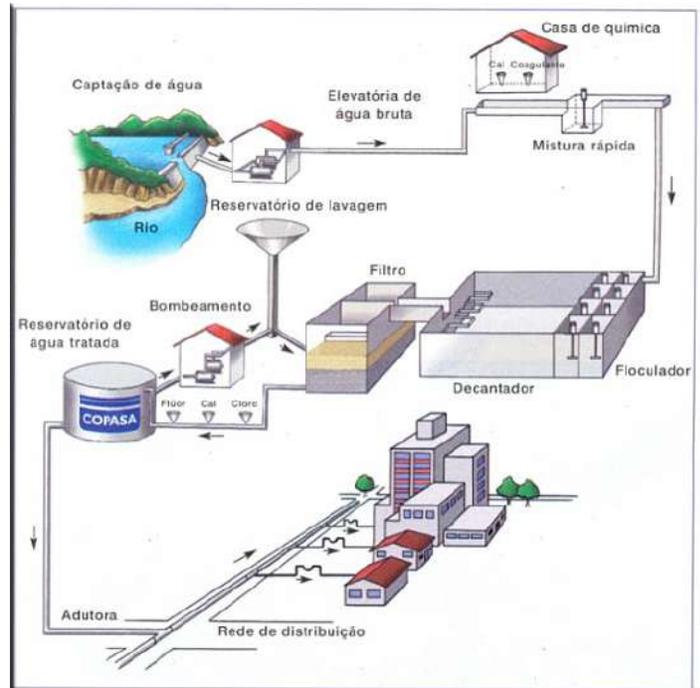


Figura 5. Exemplo de Sistema de Abastecimento.

### D. Perdas

Perdas são intrínsecas aos sistemas de abastecimento de água de tal forma, que são desconhecidos sistemas com zero

<sup>6</sup>Companhia de Saneamento de Minas Gerais.

<sup>7</sup>Tubulação para condução de água de alguma fonte específica para um reservatório.

<sup>8</sup>Sistema projetado para distribuição de água tratada.

<sup>9</sup>Interligação entre a rede de abastecimento de água e o imóvel.

de perda [5]. Apesar de ser teoricamente impossível um SAA operar sem perdas, existem restrições de natureza técnica e econômica que tornam a sua concepção e manutenção inviáveis, devido aos altos custos. Acontece que assegurar rotinas de inspeção e manutenção de todo o sistema em 100% do tempo é extremamente dispendioso, e há um limite econômico natural, onde produzir água tratada passa a ser menos custoso do que evitar as perdas [3].

Indicadores de perdas são amplamente utilizados como ferramenta gerencial para avaliação da eficiência dos SAA's e para o controle dos mananciais. Indicadores como o Índice de Perdas por Ligação (IPL) e o Índice de Perdas Totais (IPT) são utilizados pelas concessionárias prestadoras do serviço e órgãos governamentais [3].

O Sistema Nacional de Informações do Saneamento (SNIS), por exemplo, administrado pelo Ministério do Desenvolvimento Regional [10], concentra um repositório destes indicadores [11] que é amplamente divulgado no Brasil como uma ferramenta de apoio ao setor de saneamento. A Figura 6 apresenta um gráfico com o índice de perdas das principais empresas de saneamento no Brasil extraídos do Sistema Nacional de Informações do Saneamento (SNIS) [11]. Os dados se referem ao ano de 2016. Nele, pode ser visto que a empresa com maior índice de perdas CAESA<sup>10</sup>, enquanto a SANEAGO<sup>11</sup> tem o menor índice. A Embasa aparece na décima quarta posição.

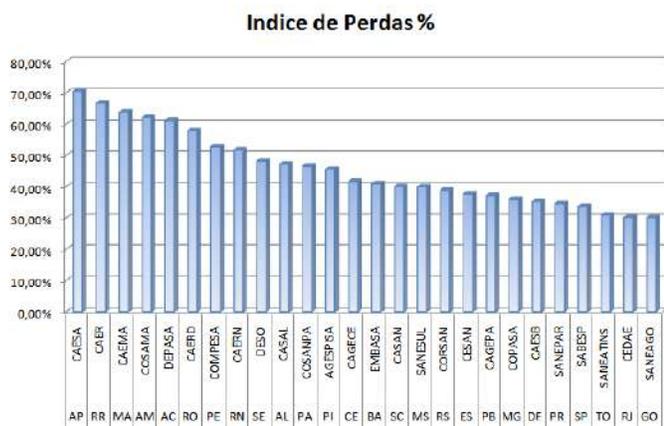


Figura 6. Comparativo entre as maiores empresas de saneamento no Brasil.

Nos últimos anos, as perdas de água no Brasil compõem números que variam entre 40% e 50%, podendo ser considerado um índice alto, sobretudo ao ser comparado com a média de países como Japão, Austrália ou Alemanha [1][2].

As perdas dos SAA's podem existir, não apenas por razões de ordem técnica, como também por outros fatores, inclusive culturais. Procede que o conhecimento de que a água é um recurso escasso e requer preservação, ainda não está completamente assimilado por nossa sociedade [5]. Um controle efetivo de perdas requer a preservação de mananciais, a melhoria da eficiência dos sistemas e a responsabilidade sobre o consumo da água, por parte da população.

Entidades governamentais e não governamentais têm demonstrado preocupação com a escassez da água. Organismos internacionais, como o *International Water Associations*<sup>12</sup> (IWA), mobilizam ações em dezenas de países, buscando medidas necessárias à redução de perdas e ao planejamento de demandas futuras [12]. No Brasil, a Associação Brasileira de Engenharia Sanitária e Ambiental (ABES) é uma das entidades que cumprem o papel de articular os setores de saneamento, compartilhando e incentivando ações voltadas à preservação dos recursos hídricos e redução de perdas [13].

Importante destacar, que o governo brasileiro vem estabelecendo diretrizes para o setor de saneamento, tendo como um dos principais objetivos a Universalização dos Serviços de Saneamento. Entenda-se universalizar como garantir o direito de todos ao acesso aos serviços de abastecimento de água, coleta de esgoto e de resíduos sólidos, e manejo de águas pluviais. Através da Lei No. 11.445, de 5 de janeiro de 2007 [14], foram definidos novos requisitos e metas que visam universalizar o acesso aos serviços de saneamento básico em médio prazo. O abastecimento de água, naturalmente, ocupa importante destaque nas exigências da lei. A universalização do abastecimento aumenta a demanda destes serviços. Assim, um melhor aproveitamento dos recursos hídricos e um controle mais efetivo de perdas tornam-se imperativos.

Além disso, diversas campanhas publicitárias, visando combater o desperdício de água, são vinculadas nas redes de comunicação em massa como mais uma forma de combate ao desperdício de água.

#### 1) Classificação das perdas:

Controle de perdas é objeto de interesse das concessionárias de abastecimento. Uma boa operação dos SAA's pode reduzir significativamente a ocorrência de vazamentos ou transbordamentos, por conseguinte os índices de perdas. Estas ações, entretanto, limitam-se à água que chega até os ramais dos consumidores, não resolvendo o problema por completo. As perdas nos sistemas não estão restritas à evasão de água da rede ou dos reservatórios, incorrendo também em problemas de outra natureza.

No setor de saneamento, existem diversas classificações para os tipos de perda, com diferentes terminologias. Todavia, os conceitos são compartilhados entre as diferentes visões [1][3], sendo a relação de consumo uma das principais classificações adotadas. Uma primeira e importante classificação das perdas consideram-nas com relação ao consumo [1][3]. A seguir, estão as duas terminologias mais difundidas entre as empresas de saneamento do Brasil:

- perdas comerciais, ou perdas aparentes [3][15], generalizam as condições onde a água tenha sido utilizada sob a forma de consumo, mas, por motivo de contabilização, não tiveram o seu consumo registrado. Incluem-se nesta classificação: submedição [15], ou outros erros de medição, ligações clandestinas e outros tipos de fraude, ligações de consumo estimado (ligações não-medidas), dentre outras;
- Perdas físicas, ou perdas reais [3], que são as perdas onde este trabalho irá atuar, envolvem os casos onde

<sup>10</sup>Companhia de Água e Esgoto do Amapá

<sup>11</sup>Saneamento de Goiás S.A.

<sup>12</sup>Associação Internacional de Águas

a água não foi utilizada sob a forma de consumo. Neste grupo, incluem-se os vazamentos nos SAA's, como, por exemplo, rompimento de tubulações ou transbordamento de reservatórios, e os seus impactos estão restritos ao controle dos mananciais e custos de operação do sistema.

Tanto as perdas físicas quanto as perdas comerciais trazem às operadoras dos sistemas um novo problema relacionado à sua receita financeira [15]. Ambas impactam negativamente no faturamento destas empresas, culminando, segundo [3], em:

- maiores dificuldades na operação dos sistemas;
- Menor disponibilidade de recursos para expansão das áreas de abastecimento;
- Penalidades financeiras à sociedade, por repercutir na composição tarifária através do reajuste de preços dos serviços.

### 2) Tecnologias implementadas:

Com relação às perdas comerciais, há tecnologias de suporte ao combate de ligações clandestinas que são, sem dúvida, objeto de interesse de todo o setor de saneamento. A incidência deste tipo de ligação é alta, e os impactos ambientais são grandes. Os prejuízos financeiros das operadoras de SAA's normalmente totalizam cerca de milhões de reais mensais.

Com relação às perdas físicas, o que está sendo feito? As pesquisas demonstram que as empresas de saneamento direcionam o maior foco de suas atenções para tentar solucionar problemas de água consumida indevidamente, já as empresas particulares, tentam solucionar problemas de vazamento intradomiciliares [16].

### 3) Cálculo de Perdas:

Este trabalho adotou, para a análise deste projeto, apenas vazamentos em tubulações chamadas ramais (ramificações de tubulação que são ligadas da rede de água para o imóvel do cliente, geralmente utilizado tubo de 1/2 polegadas) com coeficiente de descarga é de 0,62 e a coluna d'água corresponde a 20 m.c.a<sup>13</sup> (h).

A fórmula geral para pequenos orifícios [17] foi utilizada para cálculo das perdas de água. Esta equação é derivada da equação de Bernoulli<sup>14</sup> [18], que é utilizada para estudos de Hidrodinâmica. A seguir são apresentadas as equações:

$$Q = Cd * S * \sqrt{2 * g * h} \quad (1)$$

fórmula geral para pequenos orifícios, equação 1, onde:

- 'Q': Corresponde ao volume de água perdido em metros cúbicos por segundo;
- 'Cd': Coeficiente de descarga;
- 'S': Área da fissura por onde escoo o vazamento (cm<sup>2</sup>);
- 'g': Aceleração da gravidade ;

- 'h': pressão medida no local do vazamento;

$$\frac{v^2}{2} + g * h + \frac{p}{\rho} = constante \quad (2)$$

Equação de Bernoulli, equação 2, onde:

- 'v': Corresponde à velocidade do fluido ao longo do condutor;
- 'g': Aceleração da gravidade;
- 'h': Altura em relação a um referencial;
- 'p': Pressão ao longo do recipiente ;
- 'ρ': a massa específica do fluido.

## E. Ferramentas e Tecnologias

### 1) Data Base Synchronizer (DBS):

O DBS<sup>15</sup> é uma ferramenta criada para atender as demandas de integração entre bases de dados. Ele tem como objetivo realizar iterações de comandos SQL<sup>16</sup> por meio consultas SQL entre bases de dados diferentes. Esta ferramenta permite que um comando *template* seja aplicado como base para uma sequência de comandos parametrizados, por meio de outra(s) *query(ies)*<sup>17</sup> [19]. Suas operações são definidas por meio de comando SQL de atualização, que podem combinar comandos de seleção, inserção ou atualização. Estes comandos são modelados por conjuntos de dados obtidos. Por exemplo, um comando *insert*<sup>18</sup> de todos os registros da tabela *dof\_usuario* pode ser realizado na tabela *tmp\_usuario* por meio de duas chamadas declaradas conforme exemplo simplificado a seguir:

```
01 | #source = bdl.conf
02 | select matricula, nome, localidade from
   | dof_usuario;
03 |
04 | #target = bdl.conf
05 | insert into tmp_usuario(@matricula, @nome,
   | @localidade);
06 |
```

Neste exemplo, o comando ilustrado acima é executado tantas vezes quantos registros existirem no resultado da consulta declarada anteriormente no comando *select*<sup>19</sup>, coletando em cada um deles as informações dos campos *matricula*, *nome* e *localidade* [6].

### 2) Data Warehouse:

O *data warehousre* (DW) é um repositório de dados que serve para recolher informações de uma empresa para que essa possa controlar melhor um determinado processo, disponibilizando uma maior flexibilidade nas pesquisas e nas informações que necessitam [20].

O DW cria padrões, melhorando os dados analisados de todos os sistemas, corrigindo os erros e reestruturando os dados sem afetar o sistema de operação, apresentando somente um modelo final e organizado para a análise [20][21].

<sup>15</sup>Sincronizador de Banco de Dados.

<sup>16</sup>Linguagem de consulta estruturada. Ela é uma linguagem declarativa.

<sup>17</sup>Comandos utilizados para efetuar ações em banco de dados.

<sup>18</sup>Comando para inclusão de informações em banco de dados.

<sup>19</sup>Comando utilizado para recuperar informações em bancos de dados.

<sup>13</sup>Metros de coluna d'água.

<sup>14</sup>Equação utilizada para descrever o movimento de um fluido.

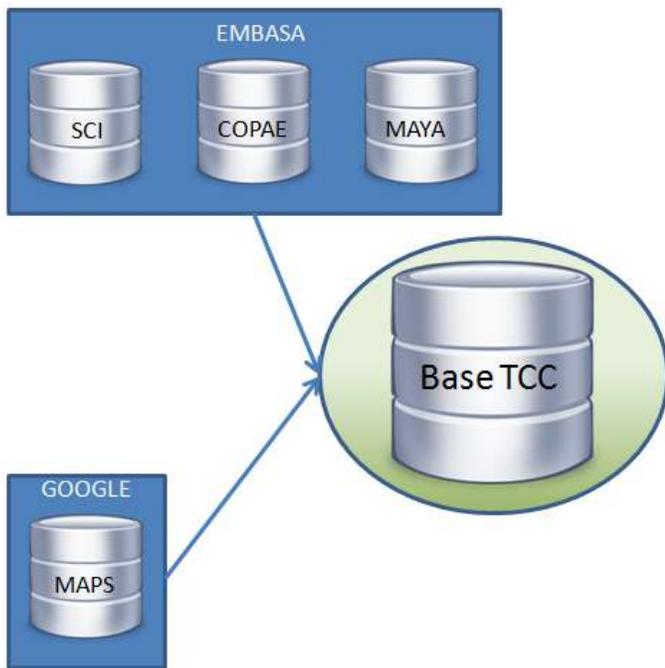


Figura 7. Processo de extração dos dados.

Com relação a criação de uma DW, é necessário efetuar uma série de diferentes passos. Estes passos correspondem aos processos de extração, transformação e carga (ETL<sup>20</sup>). Nesta etapa estão incorporados procedimentos de limpeza dos dados, integração das informações e modificação dos dados gerando uma base de dados unificada [21]. Segundo Kimball (1998) [21], o processo de ETL consome cerca de 33% do orçamento num projeto de *data warehouse* e 60% do tempo de desenvolvimento.

Ademais, é possível afirmar que a geração de um processo ETL consolida o mapeamento de todos os atributos de todas as fontes disponibilizadas para os atributos das tabelas do DW [21].

Os dados que irão compor o DW, normalmente, são disponibilizados nos mais diversos tipos de fontes (arquivos de texto, planilhas eletrônicas, bancos de dados multidimensionais, etc.). Por isso, ao iniciar o processo de extração dos dados, as informações originárias de cada fonte não devem sofrer nenhum tipo de alteração, e sim ajustes para atender ao modelo proposto do DW [20].

O DW é composto pelas tabelas dimensões<sup>21</sup> e a de fato. As dimensões são as responsáveis por armazenar os agentes qualificadores. Elas possuem descrições sobre o fato ocorrido, caracterizando o evento em si. As dimensões são reconhecidas a partir de perguntas sobre qual informação se deseja medir. A tabela fato é responsável por guardar os registros ocorridos em medições numéricas [21], sendo considerada a tabela mais importante do DW.

Existem 6 tipos de campos da tabela fato que podem ser gerados [22]:

- transacional, onde estão as métricas que podem ser somadas por todas as dimensões;
- Agregado, que é responsável por juntar um maior volume de dados que não valem a pena serem estudados no grau mais baixo;
- Consolidado, que, no geral, funciona como fato agregado mas com a vantagem de poder combinar mais de um tipo de processo (na prática equivale a firmar 2 fatos);
- *Snapshot* periódico, que se baseia em intervalos de tempo ( dia, mês, ano, etc.);
- *Snapshot* acumulado, que funciona de forma similar ao periódico, mas acumulando momentos diferentes (comumente utilizados em processos logísticos);
- Fato sem fato, quando se quer uma junção entre duas dimensões sem utilizar nenhum tipo de métrica. Ele serve para normalizar as tabelas dimensões a fim de reduzir os espaços tomados entre as tabelas, essa junção também é conhecida como esquema floco de neve.

Outro esquema bastante utilizado no processo de modelagem do DW, é o estrela. Ele pode ser considerado o modelo mais comumente empregado na modelagem dimensional para auxiliar no desempenho de sistemas que utilizam muitas consultas [21] [20].

A granularidade da tabela fato é um aspecto importante a ser observado. O grão é o grau mais baixo de informação e é determinado de acordo com as proposta exigida pelo projeto. A granularidade pode ser alta, quando possui maior detalhamento, ou granularidade baixa, quando possui menor detalhamento [23]. Ela está diretamente associada ao volume de dados armazenados, ao tempo de retorno das pesquisas e ao grau de detalhamento da DW.

Antes da criação da tabela fato, é necessário identificar quais agentes qualificadores irão compô-la, quais tabelas serão utilizadas para obter a informações necessárias, e, ao final, do processo, o esquema proposto com a uniões destas tabelas dimensões se assemelhará com o desenho de uma estrela, com tabela fato no centro e as tabelas dimensões ao redor, formando assim o esquema estrela, desenvolvido por Kimball (1998) [21].

### III. MÉTODO

O método aplicado neste trabalho consiste em três estágios:

- definição e construção de infraestrutura;
- Teste de implantação;
- Execução dos processos definidos.

O estágio de definição e construção da infraestrutura é constituído das etapas extração, transformação e carga, e visualização.

<sup>20</sup>Do inglês *Extract, Transform and Load*.

<sup>21</sup>Está no plural porque é necessário ter pelo menos 2 dimensões.

### A. Extração dos dados

Esta etapa consiste na identificação das fontes de dados, considerando:

- 1) o momento e a localização das perdas e a unidade de atendimento;
- 2) Da estrutura de armazenamento temporário das rotinas de construção e de extração.

### B. Transformação e carga dos dados

Esta etapa consiste em:

- 1) definição das dimensões tempo, endereço da solicitação, unidade organizacional e tipo de serviço;
- 2) Utilização do modelo floco de neve das dimensões que possuam uma hierarquia;
- 3) Definição dos agregados de quantidade de perdas a partir do tempo de atendimento;
- 4) Criação do modelo estrela conforme definição das dimensões e do fato;
- 5) Desenvolvimento das rotinas de transformação, considerando os dados da etapa anterior e a definição do modelo proposto. Nesta atividade, serão calculadas as perdas conforme definição.

### C. Visualização dos dados

Consiste em:

- definição da ferramenta de visualização;
- Construção da ferramenta de visualização considerando os agregados e as dimensões como filtro a serem apresentados em um mapa;
- Exibição de gráficos em linha com os dados de perdas de água e das financeiras para a localidade e por bairros;
- Exibição de gráfico radar com os dados de perdas de água e das financeiras por setor de abastecimento;
- Exibição de tabelas com informações extraídas do DW e da tabela de registros de vazamentos.

O estágio de teste de implantação consiste na criação de uma massa de dados para teste. Este estágio é composto das seguintes atividades:

- 1) Execução das rotinas definidas no estágio anterior a partir da massa de teste;
- 2) Exibição dos resultados através da ferramenta de visualização.

Apos validação de teste o próximo estágio é o de execução dos processos, que consiste na execução periódica para que haja as inclusões dos dados. Este estágio é composto das seguintes atividades:

- 1) execução das rotinas definidas na etapa de extração;
- 2) Execução das rotinas definidas na etapa de transformação e carga incluindo a validação;
- 3) Exibição dos resultados através da ferramenta de visualização.

## IV. ANÁLISE EXPLORATÓRIA

### A. Contexto Atual

A localidade de Pirajá, pertencente a Unidade Regional de Pirajá, está dividida em 21 setores de abastecimento conforme apresentado na Figura 8.

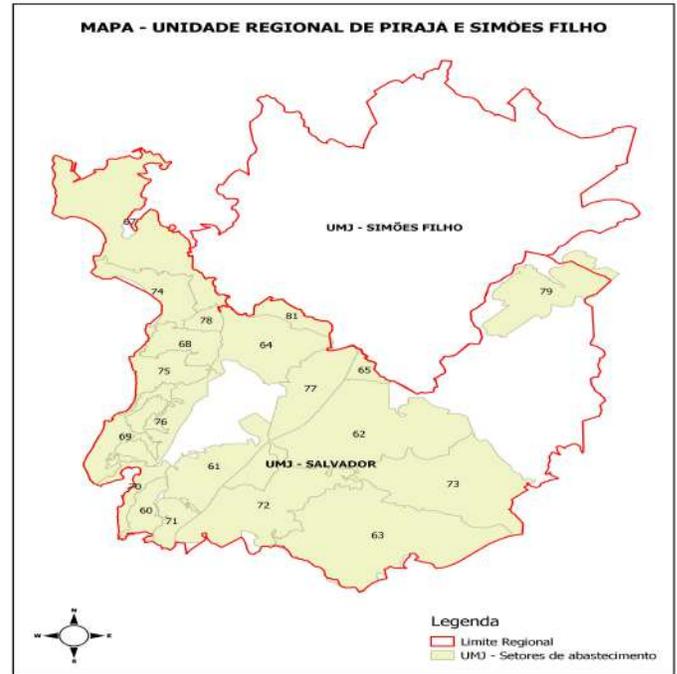


Figura 8. Setores de Abastecimento da localidade de Pirajá.

A Tabela I contém dados referentes a um dos setores de abastecimento desta localidade. Ela foi extraída do sistema Sistema de Controle Operacional de Água e Esgoto (COPAE)<sup>22</sup> [24], que só controla perdas totais.

MÊS REF.	VAZ RAMAL	VAZ REDE	VOL ESPEC	VOI FAT	VOL M_MEDIDO	VOL OFERT
JAN-2017	33	4	7980	89002	55676	117570
FEV-2017	36	3	7980	94417	55793	118460
MAR-2017	28	7	7980	93040	53796	113090
ABR-2017	28	7	7980	93939	57791	110040
MAI-2017	24	3	7980	92028	54210	105470
JUN-2017	21	1	7980	92114	53670	106320
JUL-2017	19	5	7980	93001	57887	118680
AGO-2017	25	4	7980	67459	53416	112490
SET-2017	53	27	7980	69126	55309	127970
OUT-2017	33	11	7980	70903	55152	138050
NOV-2017	27	7	7980	73214	60049	126930
DEZ-2017	28	8	7980	71365	58853	136120

Tabela I. PERDAS, SETOR 71, PIRAJÁ (COPAE).

O COPAE contabiliza a quantidade de ocorrências de vazamentos de água tanto para ramais quanto para rede. No entanto,

<sup>22</sup>Sistema responsável por armazenar dados sobre controle de distribuição e medição de água tratada e a quantidade de vazamentos que ocorreu no período.

não há aprofundamento sobre o impacto no aumento das perdas. Para efeito de cálculo, a Embasa disponibiliza dados sobre o volume de água ofertado e o que foi micromedido, mas não disponibiliza nenhum tipo de estudo que diferencie o que são perdas comerciais das perdas físicas, conforme campos descritos desta tabela:

- MÊS REF - representa o período em que ocorreu a análise;
- VAZ RAMAL - a quantidade de vazamentos em tubulações de ramais ocorridos no período;
- VAZ REDE - é a quantidade de vazamentos ocorridos em redes de água no período;
- VOL ESPEC - (Volume Especial) - é a quantidade de água disponibilizada para usos específicos (Caminhões-pipa para hospitais e instituições, utilização de corpo de bombeiros, etc.);
- VOL FAT - (Volume faturado) - é o volume de água cobrado no período;
- VOL M\_MEDIDO - (Volume micromedido), diferente do consumo faturado, este representa o volume de água efetivamente consumido no período;
- VOL OFERT - (Volume ofertado) trata-se do volume de água tratada, mensurado e disponibilizado para a zona de abastecimento no período.

De acordo com os dados apresentados nesta tabela, entre os meses de Janeiro a Dezembro de 2017, na Unidade Regional de Pirajá (UMJ), localidade de Pirajá, dentro da Zona de Abastecimento (ZA) 71, foram ofertados 136.120 metros cúbicos (m<sup>3</sup>) de água tratada em Dezembro de 2017, porém, somente 58.853 m<sup>3</sup> foram medidos (uma perda comercial de 50,90%).

Uma análise exploratória foi realizada para determinar o volume de água perdido, durante a ocorrência de vazamentos de ramais na distribuição. O valor de perdas financeiras (R\$) também foi estimado, com base na quantidade de água desperdiçada. Para tanto, é necessário a extração dos dados das seguintes fontes:

- A Embasa, que dentre outras informações, disponibilizou aproximadamente 28 milhões de registros correspondentes a todos os tipos de serviços solicitados por clientes entre os anos de 2015 e 2018. Ela determinou que fossem utilizadas apenas informações referentes ao ano de 2017, por considerar os dados mais consistentes e completos. Ao todo, ela forneceu dados específicos de três sistemas para complementar as informações:
  - 1) COPAE - com as seguintes informações: diretorias, superintendências, localidades, setores, referência, quantidade de vazamentos no período e medições de volumes (volume disponibilizado, volume micromedido, volume especial). Fornecidas através de planilhas, com dados sobre adução de água bruta distribuição e consumo de água tratada;
  - 2) SCI - com as seguintes informações: diretoria, superintendência, localidade, setores, numero

de solicitação, data hora inicio, data hora fim, tipo de vazamento, logradouro, bairro. Estes dados foram fornecidos através de visões<sup>23</sup> [25] disponibilizadas na base de dados DB2<sup>24</sup> e correspondem as informações extraídas do Sistema Comercial Integrado<sup>25</sup>;

- 3) Sistema de Gestão da Micromedição e Perdas Aparentes<sup>26</sup> (MAYA) - com as seguintes informações: município, localidades, diretorias. Estes dados foram fornecidos através de visões criadas numa base de dados Postgres.

- Google Maps - latitude e longitude. Visando obter a localização espacial dos vazamentos foram efetuadas pesquisas no Google Maps utilizando *web scraping*<sup>27</sup>

### B. Definição e Construção da Infraestrutura

No estágio de definição e construção da infraestrutura, foram analisados os dados necessários para efetuar o estudo com base nos registros de vazamentos cedidos pela empresa alvo do estudo. De posse dessas informações, foi criada uma base de dados denominada **base\_tcc** para armazenar os dados. Por convenção, ficou determinado que todas as tabelas com informações pertinentes a organização, localização, e aos serviços da empresa iniciariam o nome com o prefixo **dof\_**<sup>28</sup>, seguido de um nome simbólico significativo. Os dados relacionados a reservatórios e volumes iniciaram com prefixo **do\_**<sup>29</sup>. e os dados de localização geoespacial ficaram com prefixo **google\_**<sup>30</sup>. Estas tabelas foram criadas para atender à etapa de definição de extração dos dados. Vale ressaltar que todas estas tabelas, por padrão, foram criadas no esquema público da base de dados.

Para carga da tabela **dof\_servicos**, foi criado o artefato de extração **Perdas-SSE.sql**<sup>31</sup> que, através do uso da ferramenta DBS, efetuou sincronização entre a base de dados DB2 (utilizada pela Embasa) e o Postgres. Este artefato seleciona de uma quantidade limite de dados contidas na tabela de origem no DB2 e os insere na tabela destino **base\_tcc**. Além disso, este artefato também salva uma identificação de todos registros que foram carregado durante sua execução na tabela **dof\_processoatualizacao**. Assim, em caso de interrupção do processamento, a retomada da execução acontecerá no ponto da falha, sem haver necessidade de processamento de dados a partir do início. O código deste artefato está disponível no Apêndice A deste trabalho.

As tabelas **dof\_bairro**, **dof\_escritorio**, **dof\_especificacaoservico**, **dof\_grupofat** **dof\_localidade**, **dof\_logradouro**, **dof\_mu**

<sup>23</sup> Abreviação de uma consulta em um comando menor.

<sup>24</sup> Sistema, produzido pela IBM, que Gerencia dos Dados da Embasa.

<sup>25</sup> Sistema responsável pelo tratamento de todas as ações comerciais entre à Embasa e seus clientes, tais quais cadastros de clientes, registros de faturamento, arrecadação e serviços solicitados.

<sup>26</sup> Sistema responsável pelo controle das operações que tratam de problemas ligados a micromedição e perdas aparentes

<sup>27</sup> Técnica de extração de dados utilizada para coletar dados de sites.

<sup>28</sup> Dados oficiais obtidos através do Sistema Comercial e do Sistema Maya

<sup>29</sup> Dados operacionais disponibilizados através do Sistema COPAE

<sup>30</sup> Dados de latitude e longitude buscados via *scraping* [26]. no Google Maps

<sup>31</sup> As letras SSE significam Solicitações de Serviços Executados.

nicipio, dof\_regiao, dof\_setor, dof\_tiposolicitacao e dof\_unidaderegional foram criadas. O DBS foi utilizada para buscar dados do Sistema Maya (Postgres) e trazer os carregá-los na base\_tcc onde os dados foram armazenados criando identificação de todo registro carregados para cada uma das tabelas.

A carga das tabelas de prefixo doo\_ e google\_, foi realizada por meio de comandos de inserção dos dados contidos nas planilhas extraídas do Sistema COPAE e do Google Maps, respectivamente.

As tabelas das dimensões foram criadas consoantes descrições a seguir:

- 1) unidade Organizacional - responsável por armazenar os dados referentes às especificidades da empresa que está sendo estudada;
- 2) Endereço da Solicitação - responsável por armazenar os dados referentes à localização do serviço;
- 3) Tipo de Serviço - responsável por armazenar os dados sobre os tipos de serviço executados pela empresa;
- 4) Tempo - armazena os dados relacionados aos dias, as semanas, meses e anos.

A partir dos dados fornecidos, a proposta é a criação do Esquema Estrela (Figura 9) de DW dando importância à **perda de água**.

O modelo do DW final tem a estrutura demonstrada na Figura 9, sendo que cada dimensão ainda possui um agregado de dimensões que a compõe conforme a Figura 10.

O estudo foi realizado utilizando apenas os dados correspondentes a localidade de Pirajá. Sendo assim, os dados passaram por um processo de transformação e limpeza, e em seguida, armazenados em uma base única do Postgres (Base TCC) a fim de manter os dados padronizados e livres de inconsistências. Em seguida, a carga foi iniciada, utilizando criação das visões ou DW para o fato perdas de água, empregando o esquema estrela.

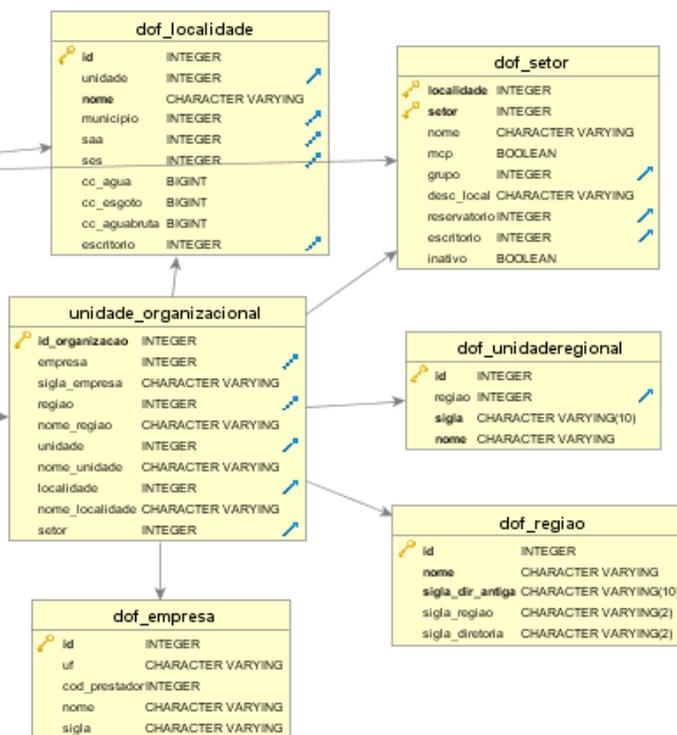


Figura 10. Dimensão Unidade Organizacional

entre tabelas pré-existentes a fim de manter em única tabela varias informações comuns entre si. Durante sua criação, foram utilizadas chaves primárias com os identificadores de cada tabela, que deixam seus itens sem necessidade de repetições, mantendo assim regra de integridade dos dados. Para separar as dimensões das outras tabelas, também foi criado um novo esquema, com nome de bi<sup>32</sup>. O Apêndice B apresenta a forma de criação de uma das dimensões deste DW.

Após conclusão desta etapa, houve uma necessidade de limpeza da tabela onde foram armazenados os registros de serviço, para manter apenas os dados relacionados ao tipo de perda por vazamentos, representado pelo código 43, com atendimento concluído. Esta etapa antecedeu à criação da tabela fato. Neste processo de limpeza, a tabela de serviços com pouco mais de 1.9 milhões de registros, sendo que:

- 547.707 eram referentes ao ano de 2017;
- 189.058 eram vazamentos de ramais
- 7.177 pertencentes à Unidade Regional de Pirajá
- 5.384 dentro da localidade de Pirajá (foco do estudo);
- 3.803 registros que possuem todos os dados de todas as dimensões (exceto Tempo) e com situação de atendidos<sup>33</sup>.

A próxima etapa culminou na criação da tabela temporária registro\_vazamentos para ser carregadas com os dados

<sup>32</sup>Do inglês *Bussiness Intelligence*.

<sup>33</sup>Como o DW está trabalhando apenas com serviços concluídos, a situação do registro atendida era essencial para a análise.

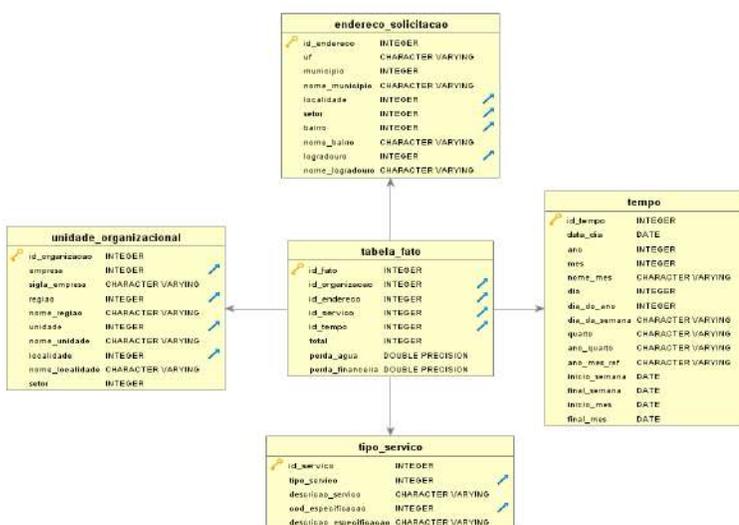


Figura 9. Esquema estrela.

Foram elaborados *scripts* para criação das tabelas que representam as dimensões. Elas utilizaram conceito de uniões

dos registros de vazamento, quebrando-os em dias, seguindo a lógica de dia da solicitação e dia do fim da execução (conclusão do conserto). Por exemplo, o serviço que iniciou dia 28/01/2017 às 16:00 e concluiu 02/02/2017 às 11:00, gerará as ocorrências para as datas 28,29,30,31/01/2017, 01 e 02/02/2017 cada qual com seu tempo devidamente calculado de acordo com o registro inicial. O Apêndice C apresenta a codificação que executa este procedimento. Na sequência, a carga da tabela de registros foi efetuada. Os registros de vazamentos foram transformados, para atender aos requisitos da dimensão tempo, de um registro em intervalo de dias para cada dia em que ele ocorreu. Nesta carga, foi feita uma junção com as dimensões para que todos os registros respeitassem as informações contidas nas dimensões de organização, endereço e serviço e assim incluir apenas registros válidos.

O *script* de criação da tabela fato contém os *id* de cada uma das tabelas dimensões. Ele é responsável por incluir um registro novo<sup>34</sup> ou adiciona os valores de perda de água e financeira a um registro já existente. Este *script* possui o mesmo padrão DBS que carrega os dados para a tabela fato e registra o *log* na tabela de relatório de carga.

A visualização dos dados foi feita com uso dos *Frameworks* Spring Boot e o AngularJS. O Spring Boot implementa a arquitetura Rest<sup>35</sup> e JPA<sup>36</sup> que torna mais simples tanto a comunicação quanto a implementação entre *backend* e *frontend* de uma aplicação. Já o AngularJS é um *framework* para implementação de *frontend* de aplicações *web*. A Figura 11 representa a visão arquitetural desta aplicação.

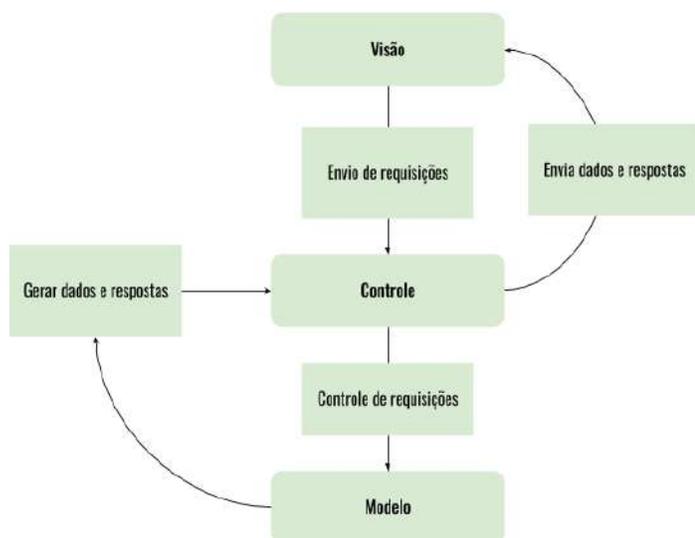


Figura 11. Visão Arquitetural da aplicação cliente/servidor.

### C. Teste da Infraestrutura

Para teste de integridade da DW, utilizou-se uma pequena massa de dados, inicialmente numa planilha eletrônica análoga à tabela fato.

<sup>34</sup>*Id* de cada uma das dimensões, total de registros, perda de água e respectiva perda financeira.

<sup>35</sup>Transferência de Estado Relacional.

<sup>36</sup>*Java Persistence API*. Para mapeamento objeto-relacional (ORM), que abstrai dados de um banco de dados para dados orientado à objetos.

Estes registros foram carregados também na tabela teste, com as mesmas características da tabela fato, no intuito de comparar os resultados da tabela de teste com a planilha eletrônica. Foram feitas consultas da tabela teste com todas as dimensões, ao mesmo tempo, e, em seguida, com cada uma separadamente como mostram os dados apresentados nos apêndices D e E respectivamente.

Numero SS	DH_Solicitacao	DH_Fim_Execucao	qtd_horas
Teste1	2017-03-07 08:29:45.608	07/03/2017 17:08	8.64
Teste2	2017-07-31 18:15:32.811	03/08/2017 11:05	64.82
Teste3	2017-12-30 13:45:15.91	01/01/2018 14:20	48.58
<b>Resultado</b>			
Teste1	07/03/2017	8,64	
Teste2	31/07/2017	5,74	
Teste2	01/08/2017	24	
Teste2	02/08/2017	24	
Teste2	03/08/2017	11,08	
Teste3	30/12/2017	10,25	
Teste3	31/12/2017	24	
Teste3	01/01/2018	14,33	

Figura 12. Massa de Teste em Planilha

```

select *
from bi.tabela_teste tt
join bi.unidade_organizacional o on tt.id_organizacao = o.id_organizacao
join bi.endereco_solicitacao e on tt.id_endereco = e.id_endereco
join bi.tipo_servico s on s.id_servico = tt.id_servico
join bi.tempo t on t.id_tempo = tt.id_tempo
  
```

id_fato_teste	id_organizacao	id_endereco	id_servico	id_tempo	total	perda_agua
integer	integer	integer	integer	integer	integer	double precision
1	1	394	25292	232	364	1
2	2	394	25292	232	365	1
3	3	394	25292	232	366	1
4	4	409	26315	232	66	1
5	5	416	23003	232	212	1
6	6	416	23003	232	213	1
7	7	416	23003	232	214	1
8	8	416	23003	232	215	1

Figura 13. Massa de Teste na DW

Vale ressaltar que os números de solicitações originais foram substituídos com o objetivo de manter a confidencialidade dos dados. O teste foi feito com registros, seguindo regras distintas:

- um registro com data de solicitação e encerramento dentro do mesmo dia, como visto no Numero SS Teste1;
- Um registro com data de solicitação no final de um mês e o encerramento no início de outro mês, como visto no Numero SS Teste2;
- Um registro com data de solicitação no final de um ano, e encerramento ano seguinte, como visto no Numero SS Teste3;

Os resultados obtidos na comparação entre a tabela teste e a planilha foram os seguintes:

O registro, denominado Teste1, que iniciou 07/03/1981 às 8h29min e concluiu 07/03/1981 às 17h08min, gerou apenas um registro na tabela fato, o registro criado, se fazer a junção com a dimensão tempo apresentou a mesma data 07/03/2017 com 8,64 horas de duração. O Teste2, que iniciou em 31/07/2017 18h15min e finalizou 03/08/2017 às 11h05min, apresentou quatro registros na tabela fato, 31/07/2017 com 5,74 horas, 01 e 02/08/2017 com 24 horas cada e 03/08/2017 com 11,08 horas, exatamente como descritos na planilha. O Teste3, que iniciou em 30/12/2017 às 13h45min e foi finalizado em 01/01/2018 às 14h20min, apresentou 3 registros na tabela fato, 30/12/2017 com 10,25 horas, 31/12/2017 com 24 horas e 01/01/2018 com 14,33 horas. Com esses testes, foi possível comprovar que o DW atende aos requisitos e pode ser utilizado efetivamente.

#### D. Estágio de Execução dos Processos

A execução do artefato de carga Perda-SSE.sql durou em torno de 36 horas. Como o processo de consulta traziam dados em intervalos de vinte mil registros, foram gerados hum mil quatrocentas linhas de *log* na tabela de processo de atualização. Os demais *scripts* foram executados na sequência, mas não levaram mais que um dia ao total para preencher as demais tabelas.

Posteriormente, iniciou-se o processo de carga das tabelas dimensões. A primeira a ser executada foi a da Unidade Organizacional. Esta durou 1 minuto e quarenta e dois segundos e gerando 3.595 registros. A dimensão seguinte foi a de Endereço da Solicitação e gerou um total de 27.716 registros. A dimensão Tipo Serviço durou menos de 6 segundos para carregar 329 registros. A dimensão Tempo, que recebeu as datas entre os anos de 2017 e 2018, totalizou 730 registros, com carga de 20 segundo.

Com as dimensões devidamente carregadas, a próxima etapa foi a carga da tabela temporária de registros de vazamentos.

Foram criadas funções<sup>37</sup> para contagem de horas de vazamentos (com base na data e hora da solicitação, e data e hora da conclusão da solicitação de serviço (SS)), e outra função específica para cálculo de perda<sup>38</sup> com base no tempo calculado na função anterior, chamado por gatilhos<sup>39</sup> [25] conforme apêndices D e E.

A junção dessas funções junto com outros processos para alimentação da tabela de registro de vazamentos está disponível no Apêndice C.

Ao chamar a função de alimentação da tabela, os registros da localidade de Pirajá se desdobraram em 6.173 linhas de registros, cada qual com a data do dia em que ocorreu, informação de duração do vazamento, quantidade de água perdida m<sup>3</sup> e quantidade de perda em reais. Vale ressaltar que, para este estudo, foi considerado Tabela Tarifária da Embasa referente ao exercício de 2017, onde a faixa mais baixa, de 0 a 6 m<sup>3</sup> de água, custa R\$ 27,50 o que dá R\$

4,58 por m<sup>3</sup> e a pressão constante em 20 m.c.a. de água<sup>40</sup>. os apêndices D e E representam os cálculos de tempo de vazamento e total de perda de água durante o período em este vazamento ocorreu. O tempo de vazamento foi calculado através de funções utilizando como parâmetros o intervalo entre início da solicitação e conclusão do serviço. O total de água perdida durante o vazamento foi calculado utilizando o resultado da fórmula geral para pequenos orifícios, jávisto, multiplicado pelo resultado obtido na função de tempo de vazamento.

A carga da tabela fato, foi executado o artefato Tabela\_-Fato.sql. A execução deste artefato segue o mesmo padrão dos anteriores, sendo que, mais uma vez, uma junção entre as dimensões é utilizada, inclusive com uso da dimensão Tempo. A seguir as principais partes do código que compõem esse artefato.

#### Junção entre as dimensões:

```
01 | join bi.unidade_organizacional dorg on dv.
    | regioao = dorg.regiao and dv.
    | unidade_regional = dorg.unidade
02 | and dv.localidade = dorg.localidade and dv.
    | setor = dorg.setor
03 | join bi.endereco_solicitacao de on dv.
    | localidade = de.localidade
04 | and dv.setor = de.setor and dv.bairro = de.
    | bairro
05 | and dv.logradouro = de.logradouro and de.
    | localidade= dorg.localidade
06 | and de.setor = dorg.setor
07 | join bi.tipo_servico ds on dv.
    | tiposolicitacao = ds.tipo_servico
08 | and dv.codespecificacao = ds.
    | cod_especificacao
09 | join bi.tempo dt on dv.datadia = dt.data_dia
```

#### Inclusão dos dados na tabela fato:

```
01 |
02 | insert into bi.tabela_fato
03 | (
04 | id_organizacao,id_endereco,id_servico,
    | id_tempo,total, perda_agua,
    | perda_financeira
05 | )
06 | values
07 | (
08 | @id_organizacao, @id_endereco, @id_servico,
    | @id_tempo, 1, @perda_m3, @perda_fin
09 | )
```

Atualização da tabela fato (em caso de já existir um registro com as chaves das quatro dimensões):

```
01 | on conflict (id_organizacao, id_endereco,
    | id_servico, id_tempo)
02 | do update set
03 | total = bi.tabela\_fato.total + 1,
04 | perda_agua = bi.tabela_fato.perda_agua+
    | @perda_m3,
05 | perda_financeira =bi.tabela_fato.
    | perda_financeira + @perda_fin;
```

<sup>37</sup>Encapsulamento de tarefas repetitivas em bancos de dados.

<sup>38</sup>Esta função representa a fórmula geral para pequenos orifícios.

<sup>39</sup>Ações específica sobre algum objeto, executado a partir de determinado evento.

<sup>40</sup>De acordo com a NBR 12.218/1994 da ABNT - Associação Brasileira de Normas Técnicas, os limites de pressão devem variar entre 10 e 50 metros de coluna d'água [27].

Geração do *log* para garantir que não haja repetição de um registro já utilizado:

```
01 | insert into bi.relatori_carga
02 | (tipo, escopo, updatedregistro, registros,
    | condicaopartida, proc_id)
03 | values
04 | (5, @filtro, now(), @#rowid, @partida,
    | @proc_id);
```

Concluído o processo, a Tabela Fato armazenou 5.736 registros, oriundos de 6173 registros.

A carga da levou pouco mais de três minutos e os dados seguiram agrupados de acordo com a união das dimensões. Os *prints* de tela com as cargas das timensões e da tabela fato estão disponíveis no D. Não foram apresentados erros durante esses processos.

### E. Novo Contexto

A ferramenta foi entregue e encontra-se em fase de homologação.

Considerando os dados contidos na tabela fato, entre os meses de Janeiro a Dezembro de 2017, na Unidade Regional de Pirajá (UMJ), localidade de Pirajá, na Zona de Abastecimento (ZA) 71 e após análises do mês de Dezembro de 2017, constata-se que foram perdidos 2.448 m<sup>3</sup> devido a vazamentos de água em ramais. Esta perda representou 3.53% do total das perdas, o equivalente a R\$ 11.212,94. Estas informações podem ser identificados na Tabela II.

A integração entre o DW e a aplicação *Web* permite identificar os locais onde ocorreram as maiores perdas de água por vazamentos de ramais.

A Figura 14, apresenta os setores de abastecimento com maior incidência de vazamentos dentro da localidade de Pirajá. As maiores ocorrências estão concentradas nas regiões do subúrbio ferroviário de Salvador e nos bairros de Pirajá, Marechal Rondon e Campinas.

Conforme dados extraídos do COPAE, a localidade de Pirajá ofertou um volume de água de 74.174.264 m<sup>3</sup> de água tratada para a população no ano de 2017. Deste volume, 46.761.726 m<sup>3</sup> foi considerado perda total. Destes 535.848m<sup>3</sup> de água foram devidos a ocorrência de vazamentos em ramais de água, equivalente 1,15% da perda total.

A Figura 15 apresenta uma das telas da aplicação onde é possível identificar os registros da tabela fato. Esses registros estão identificados por marcações de pontos dentro da região que compreende a localidade de Pirajá. Os pontos estão divididos em quatro classes:

- perdas menores que 10m<sup>3</sup>, representado por um ponto azul;
- Perdas 10m e 50<sup>3</sup>, representado por um ponto verde;
- Perdas entre 50 e 100m<sup>3</sup>, representado por um ponto amarelo;
- Perdas maiores que 100m<sup>3</sup>, representado por um ponto vermelho.

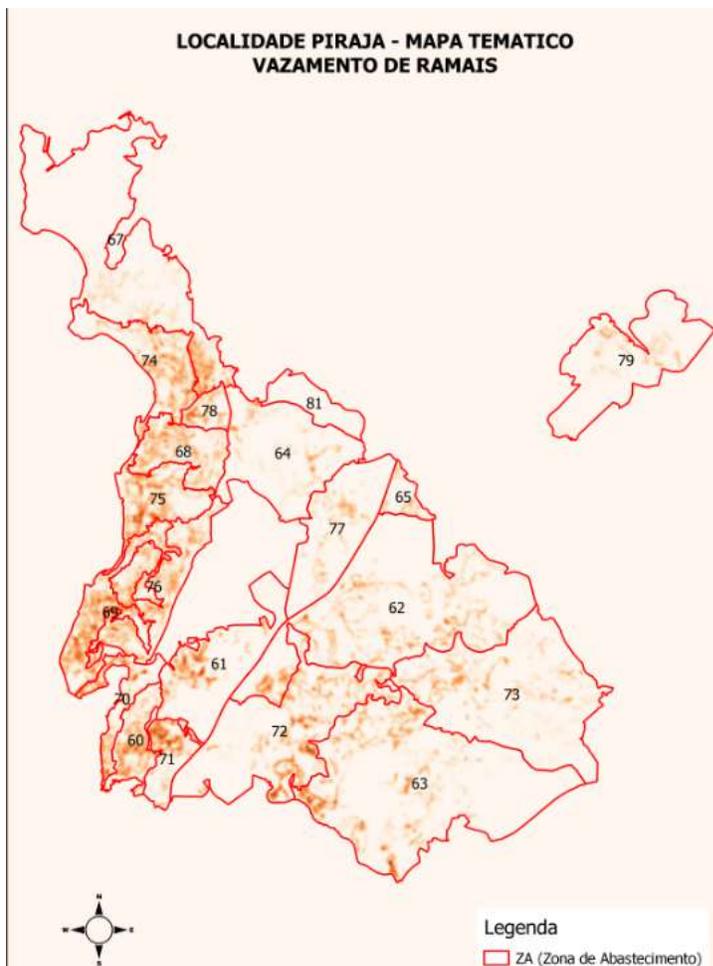


Figura 14. Mapa Térmico - Ocorrências de vazamentos.

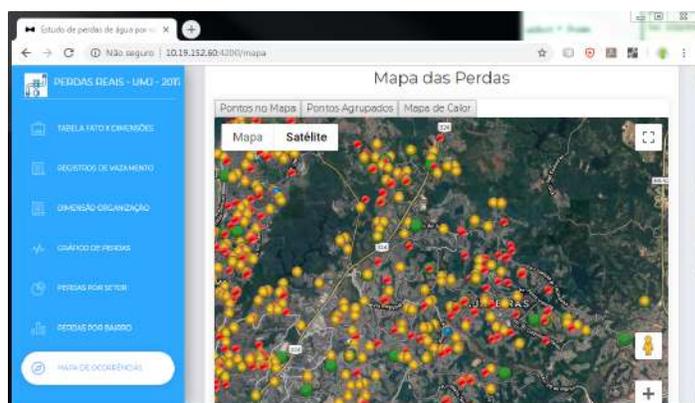


Figura 15. Mapa das perdas.

Nesta aplicação também é possível visualizar os pontos agrupado e o de mapa de calor.

A aplicação hoje está dividida da seguinte forma:

- Tabela Fato x Dimensões - contendo uma tabela unindo o fato estudado com as dimensões propostas (Unidade Organizacional, Endereço, Serviço e Tempo). As menores granularidades desta tabela são bairro,

MÊS	VAZAMENTO RAMAL	PERDA VAZAMENTO RAMAL	VOLUME OFERTADO	VOLUME MICROMEDIDO	VOLUME PERDA REAL	% VAZAMENTO RAMAL	PERDA FINANCEIRA VAZAMENTOS
JAN-2017	33	1200.78	117570	55676	53914	2.23%	R\$ 5.499,57
FEV-2017	36	3021.96	118460	55793	54687	5.53%	R\$ 13.840,58
MAR-2017	28	1203.60	113090	53796	51314	2.35%	R\$ 6.595,73
ABR-2017	28	1153.26	110040	57791	44269	2.61%	R\$ 5.281,93
MAI-2017	24	712.44	105470	54210	43280	1.65%	R\$ 3.904,17
JUN-2017	21	416.16	106320	53670	44670	0.93%	R\$ 1.907,26
JUL-2017	19	2310.72	118680	57887	52813	4.38%	R\$ 10.583,10
AGO-2017	25	2739.54	112490	53416	51094	5.36%	R\$ 12.547,09
SET-2017	53	3197.34	127970	55309	64681	4.94%	R\$ 14.643,82
OUT-2017	33	1539.96	138050	55152	74918	2.06%	R\$ 7.053,02
NOV-2017	27	1637.28	126930	60049	58901	2.78%	R\$ 7.498,74
DEZ-2017	28	2448.24	136120	58853	69287	3.53%	R\$ 11.212,94

Tabela II. CONTABILIZAÇÃO DAS PERDAS POR VAZAMENTO DE RAMAL NO SETOR 71, PIRAJÁ.

logradouro e dia. É possível exportar os dados para um arquivo texto. O cálculo da perda de água leva em consideração a pressão da água de 20 mca. Para o cálculo da perda financeira, foi levada em consideração a tabela tarifária de 2017. A tarifa na menor faixa era R\$ 27,50, ou seja, R\$ 4,58 por metro cúbico de água;

- Registros de Vazamentos - que apresenta uma tabela com os registros que foram utilizados para preencher a tabela fato. Para esta tabela, foi necessário selecionar cada solicitação e separar por dias, com o intuito de atender a dimensão tempo. A seguir um exemplo dessa separação:

Número da solicitação: XXX, início do serviço 28/01/2017 12:00, término do serviço 30/01/2017 as 9:30.

Este registro será dividido em:

XXX, 28/01/2017 | 12h00min XXX, 29/01/2017 | 24h00min XXX, 30/01/2017 | 09h30min.

Onde XXX é o número da solicitação do serviço;

- Dimensão Organização - que exibe informações sobre a dimensão Unidade Organizacional;
- Gráfico de Perdas - apresenta um gráfico comparativo das perdas de água e financeira por mês;
- Perdas por Setor - apresenta um gráfico comparativo das perdas de água e financeira por setor;
- Perdas por Bairro - apresenta um gráfico comparativo das perdas de água e financeira por bairro;
- Mapa de Ocorrências - exibe a representação espacial dos registros da tabela fato. Ele pode ser visto de 3 formas:
  - 1) Pontos no Mapa - com marcação de cada item da tabela fato no mapa. Ao clicar no ponto, ele apresenta as informações de quantidade de registros para este ponto, a quantidade de perda de água e a respectiva perda financeira;
  - 2) Pontos Agrupados - apresenta um agrupamento de registros por proximidade. Ao clicar

em um agrupamento, ele amplia o mapa da região selecionada;

- 3) Mapa de Calor - apresenta, através de mapa de calor, os locais com maior incidência de perdas.

- Sobre a Aplicação - informações para compreensão sobre a ferramenta.



Figura 16. Comparativo de perda - Água x R\$.

O gráfico de perdas gerais, representado pela Figura 16, demonstra um comparativo de perdas de água, representado pela linha amarela, e monetária representado pela linha verde, entre os meses de janeiro a dezembro de 2017, na localidade estudada.

Ainda de acordo com a Figura 16, o mês com maior incidência de perda foi o mês de julho de 2017. A quantidade de água perdida foi de 44.732,16 m<sup>3</sup> de água. Esta perda representou R\$ 204.873,96 para os cofres da Embasa.

As Figuras 17 e 18 apresentam, respectivamente, os gráficos de perdas relacionadas aos setores de abastecimento e bairros. No gráfico de perdas por setor, a linha verde representa a perda de água, já a linha azul o valor estipulado em reais. O gráfico que apresenta as perdas por bairro, utiliza o mesmo padrão de cores do gráfico de perda geral. Dentre os setores

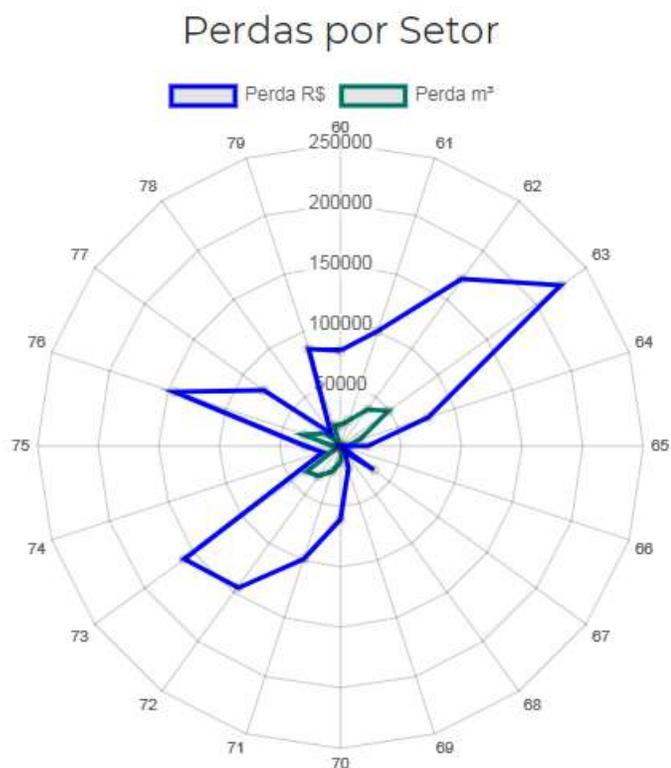


Figura 17. Perdas por setor de abastecimento.

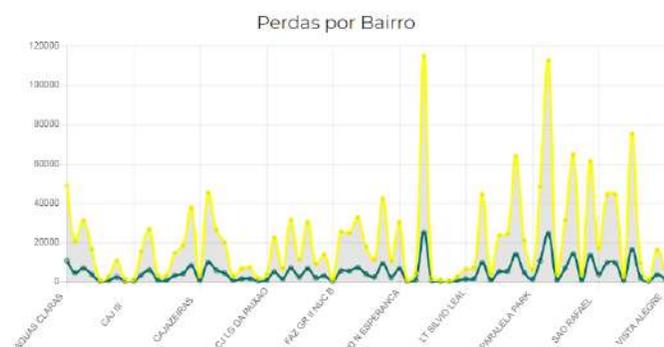


Figura 18. Perdas por bairro.

de abastecimento, a maior incidência de perdas é o setor 63, com 49.163,03 m<sup>3</sup> de água. Dentre os bairros, os que tiveram maior índice de perdas foram Lobato e Pirajá. O bairro do Lobato perdeu o equivalente a 25.130,58 m<sup>3</sup> de água, enquanto o bairro de Pirajá atingiu o índice de perda de 24.621,78 m<sup>3</sup> de água. Os respectivos bairros somados, representaram uma perda financeira de R\$ 227.867,04 do ano de 2017.

## V. CONCLUSÃO

Os objetivos do trabalho foram alcançados. Foi possível apresentar um balanço hídrico da Embasa, considerando a quantidade de água desperdiçada por vazamentos, estudando ocorrências de vazamentos de ramais em uma região dentro da cidade de Salvador. Houve uma explanação sobre o que significa balanço hídrico no contexto do saneamento, diferenciando

o que são perdas comerciais e reais. Através da elaboração do modelo, foi possível executar a análise de vazamentos ocorridos, determinar o quanto de água se perde devido a demora no atendimento das solicitações e criar visualizações, utilizando gráficos e mapas, para exibir os resultados dos dados analisados.

Como restrições, pode-se citar que não houve possibilidade de efetuar o monitoramento de solicitações em aberto até o dia anterior (**d-1**)<sup>41</sup>, visto que a solução permite carregar solicitações em andamento por agregação, na tabela de fato, de uma nova solicitação que tenha os *ids* das dimensões em outro momento, cabendo ajustar os artefatos de carga e quebra de perda por dia. Houve limitação do estudo a apenas uma pequena porção das solicitações para a realização do o trabalho, considerando o universo de registros de vazamentos ocorridos. A localidade utilizada para o estudo se concentra em apenas uma unidade, representando assim 5,26% da Embasa.

Vale ressaltar, que a perda apresentada de 535.848m<sup>3</sup> corresponde apenas a uma Unidade Regional. Considerando este valor apresentado como perda para cada uma das 19 UR's da Embasa, é possível estimar a perda total de água de 10.187.224m<sup>3</sup>. Sendo R\$ 4,58 o custo do m<sup>3</sup>, obtém-se um prejuízo próximo dos R\$ 47 milhões de reais, ao multiplicar a perda total pelo custo do m<sup>3</sup>. Este valor poderia ser investido em ações e obras para melhoria do abastecimento à população.

Na maioria dos casos, como a solicitação de serviço não sofre mudança depois de criada<sup>42</sup>, mas as outras tabelas estão sujeitas mudanças, o processo de ressetorização<sup>43</sup> fez com que alguns bairros ou logradouros mudassem de localidade e setor, impactando na eliminação de quase 1500 registros de vazamentos que não atendiam às dimensões propostas do DW. É importante destacar, que há obrigatoriedade no registro de todas as solicitações, melhor dizendo, os funcionários são orientados a realizar o registro das solicitações com dados incompletos e consequentemente com campos não preenchidos (principalmente endereços e referências), o que precisou ser filtrado para atender as dimensões.

A perspectiva para esse trabalho é a implementação da ferramenta criada e sua integração com o sistema comercial da empresa em todas as localidades, regiões e diretorias, bem como para difundir junto às demais empresas de saneamento.

Como trabalhos futuros, esperamos que com a disponibilização das tabelas de pressão por região, será possível determinar valores, mais precisos a respeito da perda de água. Também está nas pretensões, desenvolver uma arquitetura para abarcar o acompanhamento de desempenho das equipes que atendem as solicitações.

É importante destacar, que os benefícios deste trabalho vão muito além da identificação, redução do tempo entre a descoberta de vazamento de água, conserto e economia da empresa com as perdas oriundas de vazamento. Essa pesquisa também teve a preocupação em fazer essa abordagem para encontrar mecanismos de cessar e/ou minimizar o desperdício de água por se tratar de um recurso natural valioso. É possível e indispensável promover a sustentabilidade do nosso planeta.

<sup>41</sup>Identificação de dia anterior, ou seja, dia-1.

<sup>42</sup>Podendo apenas atualizar observações, cancelar ou encerrar a solicitação.

<sup>43</sup>Mudança de setores de abastecimento de localidade ou setor de abastecimento.

## APÊNDICE A EXEMPLO DE FUNÇÃO UTILIZANDO O DBS

O código abaixo foi utilizado para carregar as informações da base de dados da Embasa (Banco DB2), para a base de dados do TCC (Base TCC).

Este valor representa o título do processo que ficará sendo apresentado na tela.

```
01 | #macro = SSE
02 | $$SHOW_SCENE=false
03 | $$SHOW_STATUS=false
```

Esta parte do código representa o arquivo de configuração que o programa DBS utiliza para fazer conexão na base de dados Base\_TCC.

```
01 | #source = db1.conf
```

Nesta parte do código, é gerado um número randômico que será definido como id do processo que estará sendo executado.

```
01 | select random(99999999) as proc_id
```

Nesta etapa, é criada a condição de partida do processo de leitura, na primeira parte, o código confirma se a data de atualização não é nula. Esta informação é salva como filtro. O campo @\$MULTIOP faz uma comparação entre os dados. Primeiro ele testa se o primeiro campo é >= ao equivalente com sinal de @, caso atenda ele faz a segunda comparação. Esse texto é salvo como partida.

```
01 | #source = db1.conf
02 | select 'SSETMATU IS NOT NULL' as filtro,
03 |     '@$MULTIOP{ (SSETMATU,
           SSENNUMSOL) >= ('@DATA_ATUALIZACAO::
           native',
           @NUMEROSOLICITACAO) }' as partida
04 | ;
```

Esta parte do código representa o arquivo de configuração que o programa DBS utiliza para fazer conexão na base de dados do DB2 da Embasa.

```
01 | #source = db2-4200-H.conf
02 | $REPEAT_WHILE_FOUND=true
```

O código a seguir, representa uma seleção na base de dados da Embasa.

```
01 | select
02 | SSENNUMSOL AS numerosolicitacao,
03 | SSECDTPS AS tiposolicitacao,
04 | SSECDESP AS codespecificacao,
05 | SSECDMSO AS cod_motivo_solicitacao,
06 | SSECDSTSOL AS situacao_solicitacao,
07 | SSEICATENDIDA AS indicador_atendida,
08 | SSECDRAT AS codigo_rateio,
09 | SSECDTHRSOL AS datahroasolicitacao,
10 | SSECDTHRFIMEXEC AS datahorafimexecucao,
11 | SSECDTHRCONCL AS datahoraconclusao,
12 | SSECDUNT AS cod_unidade_anterior,
13 | SSECDUNDSOL AS cod_unidade_solicitante,
```

```
14 | SSECDUNDEST AS unidade_destino,
15 | SSECDUNDATUAL AS unidade_atual,
16 | SSENMMATUSU AS matricula,
17 | SSECDLOC AS localidade,
18 | SSECDSFT AS setor,
19 | SSENNUADRA AS quadra,
20 | SSECDBAIRRO AS bairro,
21 | SSECDLOG AS logradouro,
22 | SSENIMOVEL AS porta,
23 | SSENCEP AS cep,
24 | SSECDMUN AS municipio,
25 | SSECDTPA AS tipoarea,
26 | SSECDDBES AS cod_esgoto_sanitario,
27 | SSECDLCR AS cod_local_ocorrendia,
28 | SSECDPRU AS pavimento_rua,
29 | SSECDPCA AS pavimento_calçada,
30 | SSECDTPREVINIC AS data_previsao_inicial,
31 | SSECDTPREVATU AS data_previsao_atual,
32 | SSECDMRA AS motivo_reabertura_ss,
33 | SSECDSMC AS motivo_cancelamento,
34 | SSENNUMSOLVINC AS numero_solic_viculada,
35 | SSEDCOMPL AS complemento_endereco,
36 | SSEDSPTOREF AS referencia,
37 | SSEDSOBS AS observacao,
38 | SSETMATU AS data_atualizacao,
39 | SSECDSUP AS unidade_regional,
40 | SSECDPAD AS polo,
41 | SSENNPRAZO AS prazo_atendimento,
42 | SSECDUNTATUAL AS unidade_solicitacao
43 | from BIADM.SCITSSSE
```

Nesta parte do código, ele consulta a informação inserida na variável filtro. O código `::native` é utilizado para remover as aspas inseridas na variável.

```
01 | where @filtro::native
```

Nesta parte ele consulta a informação salva no campo partida. O comando `@query=db1.conf` é responsável por fazer uma sub-consulta dentro da consulta. Aqui também é definido o tipo de processo de atualização e limitado a vinte mil linhas por `loop`.

```
01 | @query=db1.conf{select ' and (||
           condicaopartida||)' from
           dof_processoatualizacao
02 | where tipo = 9 and escopo = @filtro order by
           id desc limit 1 }::native
03 | and SSETMATU > '2018-08-20 18:18:33.193879'
04 | order by SSETMATU
05 | fetch first 20000 rows only
```

Este campo é responsável por efetuar o `loop`, enquanto tiver informação na base de dados de origem.

```
01 | #iffound=10
02 | repeat;
```

Caso não encontre ele para a execução com a instrução abaixo.

```
01 | #ifnotfound
02 | stop;
```

O campo `target` aponta para o arquivo de configuração de entrada, para a base de dados de destino.

```
01 | #target = db1.conf
```

O campo a seguir, define o tamanho do textitbatch de inserção na tabela de destino.

```
01 | $BATCH_SIZE=20000
02 | $SHOW_SQL=false
```

O campo acima, determina se deve ser mostrado na tela os códigos resultantes.

A linha de comando abaixo segue a regra de inserção da coluna. Ela equivale a texto "para a linha X, inserido NUMEROSOLICITACAO= XYZ SSETMATU= DD/MM/AAAA HH:MM".

```
01 | $LABEL=INSERT \#rowid: NUMEROSOLICITACAO=
    @numerosolicitacao SSETMATU=
    @data_atualizacao
```

Nesta etapa ocorre o inserção dos dados na tabela de destino, de acordo com os dados de origem. O sinal @\* representa os dados obtidos no processo de seleção dos dados na base da Embasa.

```
01 | insert into dof_servicos ( numerosolicitacao
    ,tiposolicitacao, codespecificacao,
    codtipoorigemsolicitacao,
02 | codmotivosolicitacao, situacaosolicitacao,
    indicadoratendida,
03 | codigorateio,
04 | datahorasolicitacao,
05 | datahorafimexecucao, datahoraconclusao,
    codunidadeanterior, codunidadesolicitante
    ,
06 | unidatedestino, unidateatual, matricula,
    localidade, setor, quadra, bairro,
07 | logradouro,
08 | porta, cep, municipio, tipoarea,
    codesgotosanitario,
09 | codlocalocorrenca, pavimentorua,
10 | pavimentocalcada, dataprevisaoinicial,
    dataprevisaoatual,
    mat_funcionario_atualizou,
11 | complemento_endereco, ponto_referencia,
    obsevacao, num_solic_reativada,
12 | funcionario_registrou, data_atualizacao,
    data_hora_registro,
13 | media_atendimento, regioa, unidade_regional,
    prazo_atendimento,
    tipo_abertura_especifica,
14 | unidade_atua_solicitacao)
15 | values
16 | (
17 |     @*
18 | )
```

Após inserção dos dados, ele insere o log na tabela dof\_processoatualizacao com informações referentes ao processo obtido. Esses dados serão consultados a partir do segundo loop de seleção em diante.

```
01 | #afterlast = db1.conf
02 | \$$RECURSIVE_REFERENCE=true
03 | insert into dof_processoatualizacao (tipo,
    escopo, updatedregistro, registros,
    condicaopartida, proc_id)
```

```
04 | values (9, @filtro, @DATA_ATUALIZACAO, @#
    rowid, @partida, @proc_id);
```

## APÊNDICE B CRIAÇÃO DE UMA DAS DIMENSÕES DO DW

Comando para criação da tabela tipo\_servico, uma das dimensões do DW.

```
01 | CREATE TABLE bi.tipo_servico
02 | (
03 |     id_servico integer NOT NULL DEFAULT
    nextval('bi.servico_id_seq'::regclass),
04 |     tipo_servico integer,
05 |     descricao_servico character varying,
06 |     cod_especificacao integer,
07 |     descricao_especificacao character varying,
08 |     CONSTRAINT servico_pkey PRIMARY KEY (
    id_servico),
09 |     CONSTRAINT fk_serv_especificacao FOREIGN
    KEY (tipo_servico, cod_especificacao)
10 |     REFERENCES public.
    dof_especificacaoservico (codsolicitacao
    , codespecificacao) MATCH SIMPLE
11 |     ON UPDATE NO ACTION ON DELETE NO
    ACTION,
12 |     CONSTRAINT fk_serv_tiposervico FOREIGN KEY
    (tipo_servico)
13 |     REFERENCES public.dof_tiposolicitacao
    (id) MATCH SIMPLE
14 |     ON UPDATE NO ACTION ON DELETE NO
    ACTION,
15 |     CONSTRAINT
    tipo_servico_solicitacao_especificacao_key
    UNIQUE (tipo_servico, cod_especificacao
    )
16 | );
```

## APÊNDICE C FUNÇÃO DE SEPARAÇÃO DE REGISTROS POR DIA

Função responsável por separar uma solicitação única, durante um intervalo de dias em cada dia em que a solicitação ocorreu.

```
01 | CREATE OR REPLACE FUNCTION bi.
    horas_dia_vazamento(
02 |     nss integer,
03 |     ts integer,
04 |     ce integer,
05 |     reg integer,
06 |     ur integer,
07 |     loc integer,
08 |     setor integer,
09 |     bai integer,
10 |     logr integer,
11 |     mun integer,
12 |     dhsol timestamp without time zone,
13 |     dhexec timestamp without time zone)
14 |
15 | RETURNS void AS
16 | $BODY$
17 |
18 | DECLARE
19 |
20 | ldata_ini date;
21 | ldata_fim date;
22 | lhora time;
23 | lhora_ini time;
24 | lhora_fim time;
```

```

25 | 124hs integer;
26 | i integer;
27 | r_lat varchar;
28 | r_long varchar;
29 |
30 | BEGIN
31 | select trunc(round((EXTRACT(epoch FROM(
    datahorafimexecucao -
    datahorasolicitacao))/3600::double
    precision)::numeric, 2)/24)
32 | into 124hs from dof_servicos s where s.
    numerosolicitacao = nss;
33 |
34 | SELECT date(datahorasolicitacao) from
    dof_servicos s where s.numerosolicitacao
    = nss into ldata_ini;
35 |
36 | SELECT date(datahorafimexecucao) from
    dof_servicos s where s.numerosolicitacao
    = nss into ldata_fim;
37 |
38 | SELECT (datahorasolicitacao::time) from
    dof_servicos s where s.numerosolicitacao
    = nss into lhora;
39 |
40 | SELECT ('24:00:00' - datahorasolicitacao::
    time) from dof_servicos s where s.
    numerosolicitacao = nss into lhora_ini;
41 |
42 | SELECT datahorafimexecucao::time from
    dof_servicos s where s.numerosolicitacao
    = nss into lhora_fim;
43 |
44 | if (lhora > lhora_fim)
45 | then
46 |     124hs = 124hs - 1;
47 | else
48 |     124hs = 124hs - 2;
49 | end if;
50 |
51 | --dia inicial ou se o servico iniciou e
    terminou no mesmo dia
52 |
53 | insert into bi.registro_vazamento (
    numerosolicitacao, tiposolicitacao,
    codespecificacao, regiao,
    unidade_regional, localidade, setor,
    bairro, logradouro, municipio,
    datahorasolicitacao,
    datahorafimexecucao, datadia,
    qtd_horas_dia, qtd_perda_dia,
    valor_perda, latitude, longitude)
54 | values
55 | (nss, ts, ce, reg, ur, loc, setor, bai,
    logr, mun, dhsol, dhexec,
    ldata_ini,
56 | (select round((EXTRACT(epoch FROM(
    lhora_ini)/3600::double precision))::
    numeric, 2)::double precision),
57 | (select round((((0.61 * ((0.018/2)^2 * pi
    ())/2::numeric)::double precision) * /
    (2*9.8 * 20::numeric)::double precision)
    )
58 | * 3600::double precision)*round((EXTRACT(
    epoch FROM(lhora_ini)/3600::double
    precision)::numeric, 2)::double
    precision) ,
59 | (select round((((0.61 * ((0.018/2)^2 *
    pi ())/2::numeric)::double precision) *
    (|/ (2*9.8 * 20::numeric)::double
    precision))
60 | * 3600::double precision)*round(round((
    EXTRACT(epoch FROM(lhora_ini)/3600::
    double precision)::numeric, 2) * 4.58::
    numeric, 2)::double precision),

```

```

61 | (select t.latitude from tmp_servicos_cep
    t where t.localidade = loc and t.
    logradouro = logr limit 1),
62 | (select t.longitude from tmp_servicos_cep
    t where t.localidade = loc and t.
    logradouro = logr limit 1) );
63 |
64 | --inicio do loop para intervalo maior que 2
    dias
65 |
66 | i=0;
67 |
68 | while (i <= 124hs) loop
69 |
70 |     i = i + 1;
71 |
72 | insert into bi.registro_vazamento (
    numerosolicitacao, tiposolicitacao,
    codespecificacao, regiao,
    unidade_regional, localidade, setor,
    bairro, logradouro, municipio,
    datahorasolicitacao, datahorafimexecucao
    , datadia, qtd_horas_dia, qtd_perda_dia,
    valor_perda, latitude, longitude)
73 |
74 | values
75 | (nss, ts, ce, reg, ur, loc, setor, bai,
    logr, mun, dhsol, dhexec,
    ldata_ini+i,
76 | (select round((EXTRACT(epoch FROM(time'
    24:00:00')/3600::double precision))::
    numeric, 2)::double precision),
77 | (select round((((0.61 * ((0.018/2)^2 * pi
    ())/2::numeric)::double precision) * (|/
    (2*9.8 * 20::numeric)::double precision
    ))
78 | * 3600::double precision)*round((EXTRACT(
    epoch FROM(time'24:00:00')/3600::double
    precision)::numeric, 2)::double
    precision) ,
79 | (select round((((0.61 * ((0.018/2)^2 *
    pi ())/2::numeric)::double precision) *
    (|/ (2*9.8 * 20::numeric)::double
    precision))
80 | * 3600::double precision)*round(round((
    EXTRACT(epoch FROM(time'24:00:00')
    /3600::double precision)::numeric, 2) *
    4.58::numeric, 2)::double precision),
81 | (select t.latitude from tmp_servicos_cep
    t where t.localidade = loc and t.
    logradouro = logr limit 1),
82 | (select t.longitude from tmp_servicos_cep
    t where t.localidade = loc and t.
    logradouro = logr limit 1) );
83 | end loop;
84 |
85 | -- se tiver 2 ou mais dias, essa funcao e
    chamada para calcular o dia final
86 |
87 | if (ldata_fim > ldata_ini) then
88 |
89 | insert into bi.registro_vazamento (
    numerosolicitacao, tiposolicitacao,
    codespecificacao, regiao,
    unidade_regional, localidade, setor,
    bairro, logradouro, municipio,
    datahorasolicitacao, datahorafimexecucao
    , datadia, qtd_horas_dia, qtd_perda_dia,
    valor_perda, latitude, longitude)
90 |
91 |
92 | values
93 |
94 | (nss, ts, ce, reg, ur, loc, setor, bai,
    logr, mun, dhsol, dhexec,
    ldata_fim,
95 | (select round((EXTRACT(epoch FROM(

```

```

        lhora_fim)/3600::double precision))::
        numeric, 2)::double precision),
96 | (select round(((0.61 * ((0.018/2)^2 * pi
    )/2)::numeric)::double precision) * (|/
    (2*9.8 * 20::numeric)::double precision
    ))
97 | * 3600::double precision)*round((EXTRACT(
    epoch FROM(lhora_fim)/3600::double
    precision)::numeric, 2)::double
    precision ),
98 | (select round(((0.61 * ((0.018/2)^2 *
    pi()/2)::numeric)::double precision) *
    (|/ (2*9.8 * 20::numeric)::double
    precision))
99 | * 3600::double precision)*round(round((
    EXTRACT(epoch FROM(lhora_fim)/3600::
    double precision))::numeric, 2)* 4.58::
    numeric, 2)::double precision),
100 | (select t.latitude from tmp_servicos_cep
    t where t.localidade = loc and t.
    logradouro = logr limit 1),
101 | (select t.longitude from tmp_servicos_cep
    t where t.localidade = loc and t.
    logradouro = logr limit 1) );
102 |
103 | end if;
104 |
105 | END$BODY$
106 |
107 | LANGUAGE plpgsql VOLATILE
108 |
109 | COST 100;
110 |
111 | ALTER FUNCTION bi.horas_dia_vazamento(
    integer, integer, integer, integer,
    integer, integer, integer, integer,
    integer, integer, timestamp without
    time zone, timestamp without time zone)
112 |
113 | OWNER TO postgres;
114 |
115 |

```

#### APÊNDICE D FUNÇÃO PARA CÁLCULO DE TEMPO EM QUE HOUE VAZAMENTO

```

01 |
02 | CREATE OR REPLACE FUNCTION public.
    calculo_horas(nss integer)
03 | RETURNS double precision AS
04 | $BODY$
05 | DECLARE
06 | lhoras float;
07 | BEGIN
08 |
09 | select round((EXTRACT(epoch FROM(
    datahorafimexecucao -
    datahorasolicitacao))
10 | /3600::double precision)::numeric,2) into
    lhoras
11 | from dof_servicos
12 | where numerosolicitacao = nss;
13 | --limit 1;
14 | if lhoras < 0 then
15 | return 0;
16 | end if;
17 | return lhoras;
18 | END$BODY$
19 | LANGUAGE plpgsql STABLE
20 | COST 100;
21 | ALTER FUNCTION public.calculo_horas(integer)
22 | OWNER TO postgres;

```

```

23 |
24 |

```

#### APÊNDICE E FUNÇÃO PARA CÁLCULO DO TOTAL DE PERDA DE ÁGUA NO VAZAMENTO

```

01 | CREATE OR REPLACE FUNCTION public.
    calculo_perda(nss integer)
02 | RETURNS double precision AS
03 | $BODY$
04 | DECLARE
05 | lperdas float;
06 | BEGIN
07 | SELECT round(((0.61 * ((0.018/2)^2 *
    pi()/2)::numeric)::double precision
08 | * (|/ (2*9.8 * 20::numeric)::double
    precision))
09 | * 3600::double precision)* calculo_horas(
    nss)::numeric, 2) into lperdas
10 | FROM view_vazamentos_atendidos
11 | WHERE numerosolicitacao = nss;
12 | --limit 1;
13 | if lperdas < 0 then
14 | return 0;
15 | end if;
16 | return lperdas;
17 | END$BODY$
18 | LANGUAGE plpgsql STABLE
19 | COST 100;
20 | ALTER FUNCTION public.calculo_perda(integer)
21 | OWNER TO postgres;
22 |

```

Salvador-BA, 21 de ~~Fev~~ de 2019.

Ao Senhor  
**Paulo Jenner Santos Borges**  
Gerente da Unidade de Redução e Controle de Perdas - TRP

Assunto: **Autorização para uso de dados.**

Senhora Gerente,

Solicito autorização para utilização de dados da Embasa, no que se refere ao assunto de perdas de água, com a necessária permissão para consulta a dados e afins.

Esclareço que a pesquisa trata de um estudo de perdas de águas decorrentes de vazamentos de ramais, para fins de elaboração de **Trabalho de Conclusão de Curso do Instituto Federal da Bahia**.

Vale ressaltar que o solicitante é funcionário dessa empresa (matrícula 12778-7, contratado em Novembro de 2006), já trabalha com projeto de perdas na gerência onde está lotado e está ciente sobre as normas da empresa e sobre a necessidade de sigilo sobre determinadas informações.

Atualmente está trabalhando com dados disponibilizados no Sistema Nacional de Informações sobre Saneamento (SNIS), COPAE e dados de registros de vazamentos de ramais.

O foco de estudo será a Unidade Regional de Pirajá. O período será entre Janeiro e Dezembro de 2017. Serão utilizados dados referentes a solicitações de conserto de vazamentos de ramais, utilizando-se apenas dados de localização (localidade, setor, bairro, logradouro), serviços (tipo e especificação), data/hora (solicitação do serviço, execução do serviço).

A partir destes, serão determinados o tempo de vazamento, e estimados quantidade de água perdida por vazamento (utilizando conceitos de Mecânica dos fluidos).

Segue anexo TCC (ainda em fase de produção) para conhecer.

Respeitosamente,

Ricardo Martins Carvalho  
TRPA - Gerência de Micromedição e Perdas Aparentes  
EMAIL: ricardo.martins@embasa.ba.gov.br  
IFBA - Instituto Federal e Tecnológico da Bahia  
Curso - Tecnologia em Análise e Desenvolvimento de Sistemas  
ricardo.carvalho@ifba.edu.br  
Telefone: (71) 99251-5967





INSTITUTO FEDERAL DE  
EDUCAÇÃO, CIÊNCIA E TECNOLOGIA  
BAHIA  
Campus Salvador

Ministério da Educação  
Secretaria de Educação Profissional  
e Tecnológica

Salvador, 25 de Fevereiro de 2019

À Empresa Baiana de Águas e Saneamento S.A

01872

2019

*Leonel A. Santos*  
Mat.: 9612-1  
GPEC - Prot.

O aluno Ricardo Martins Carvalho, estudante de graduação em Tecnologia em Análise e Desenvolvimento de Sistemas no Instituto Federal da Bahia, matrícula 2016116048, e Assistente de Saneamento na Embasa, matrícula 12778-7, lotado na Gerência de Micromedição e Perdas Aparentes (TRPA), proprietário do email ricardo.martins@embasa.ba.gov.br e atende no ramal de telefone 3373-7587, está matriculado na disciplina Trabalho Final de Curso, código INF023. Para realização do seu Trabalho de conclusão, o aluno e seu orientador Jowaner de Oliveira Araújo (Analista de Sistemas) escolheram como tema de seu trabalho, estudo de perdas reais de água, com foco em vazamentos de água. Este estudo também despertou interesse da Unidade (Unidade de Redução e Controle de Perdas) cuja qual o setor ao qual o estudante/funcionário está ligado. Para que seu TCC seja viável, solicitamos autorização de pesquisa conforme orientação da Unidade de Gestão do Conhecimento, para utilizar o histórico de informações referente a solicitações de serviços relacionadas a vazamentos de água para uma análise estatística de cunho acadêmico.

Atenciosamente,

*Jowaner Araújo*

Prof. Me. Jowaner de Oliveira Araújo

Orientador

jowaner@ifba.gov.br

71 - 21029489

MAA 026 8787



A DT,

Favor deliberar sobre o pedido apresentado no anexo, temos interesse e estamos acompanhando o trabalho desenvolvido.

Em 21/02/19

Paulo Jenner Santos Borges

Paulo Jenner Santos Borges

TRP - Unidade de Redução e Controle de Perdas  
Mat.: 10.027-7

EMBRASA  
Diretoria Técnica e de Planejamento - DT

Em. 25.02.19

Ass. Autos

A TRP:

De ordem da DT,

Solicitamos que para o atendimento ao pleito, e necessário o envio de Ofício da Justiça, em papel timbrado, explicitando o objeto da pesquisa, abra

para todos os envolvidos.

25/02/19

Terezinha

Terezinha Loliola da Cruz Souza  
Assessora da Diretoria Técnica e Planejamento  
Mat.: 7019-A - A-DT

A TRPA

Favor solicitar ao interessado a inclusão do documento solicitado.

Em 08/03/2019

Paulo Jenner Santos Borges

Paulo Jenner Santos Borges

TRP - Unidade de Redução e Controle de Perdas  
Mat.: 10.027-7

A A-DT

Conforme entendimentos mantido via telefone com o Sr. Antonio, estamos devolvendo essa documentação para ser anexada ao mesmo solicitado.

Em: 08/03/19

Fabrizio Mota Oliveira  
Gerente de Micromedicação  
TRPA



ANEXO B  
DETERMINAÇÃO PARA CUMPRIMENTO DA SOLICITAÇÃO

19/03/2019

Gmail - Atendimento ao Pleito GPSiS 7047



Fabricio Mota <fabricio.mota@gmail.com>

---

**Atendimento ao Pleito GPSiS 7047**

2 mensagens

---

**Diretoria Técnica e de Planejamento - DT** <dt@embasa.ba.gov.br> 15 de março de 2019 15:41  
Para: FABRICIO MOTA OLIVEIRA - TRPA <fabricio.oliveira@embasa.ba.gov.br>  
Cc: TEREZINHA LOIOLA DA CRUZ SOUZA - A-DT <terezinha.loiola@embasa.ba.gov.br>, JOAO AUGUSTO SOARES SEIXAS - TRP <joao.seixas@embasa.ba.gov.br>

Fabrcio,

De ordem, pedimos providenciar atendimento ao pleito 7047, cadastrado no Sistema GPSiS.

Já solicitamos à DP a inclusão dos gestores da TRP. Enviamos esta cópia por email para que o atendimento seja iniciado.

Grata,

Lidiane I. Alabi  
Secretária da Diretoria Técnica e de Planejamento - DT  
Tel.: 71 3372-4230



[www.embasa.ba.gov.br](http://www.embasa.ba.gov.br)

As informações deste documento e todos os demais a este anexado são confidenciais e legalmente protegidos, somente podendo ser usados pelo indivíduo ou entidade a quem foi endereçado e exclusivamente para os fins institucionais para os quais se propõem. A utilização, divulgação, disponibilização, ou qualquer forma de compartilhamento destas informações ficam restritas às pessoas para as quais foram endereçadas e exclusivamente para os fins institucionais para os quais foram produzidos sendo ilícita qualquer veiculação ou utilização fora destes parâmetros e sujeitando o responsável a sanções disciplinares administrativas e penalmente previstas, tais como disciplinadas pelos art 482 da CLT e art 154 do Código Penal.

*The information contained in this document and all its attachments are confidential and legally protected, only for use by the intended recipient and exclusively for the institutional purposes for which it was made. The use, disclosure, release, or any form of sharing of this information is restricted to the receivers for which it was addressed and exclusively for institutional purposes for which they were produced, becoming illegal all forms of broadcasting or use outside these parameters and subjecting its responsible to administrative and disciplinary sanctions and criminal penalties, such as regulated by Article 482 of the Brazilian Labor code and by Article 154 of the Brazilian Criminal Code.*

---

 **OF IFBA 25 02 19 Discente Ricardo Carvalho.pdf**  
1458K

---

**Diretoria Técnica e de Planejamento - DT** <dt@embasa.ba.gov.br> 15 de março de 2019 15:49  
Para: FABRICIO MOTA OLIVEIRA - TRPA <fabricio.oliveira@embasa.ba.gov.br>  
Cc: TANIA CANDILAURO VELOSO FRANCA - TRPA <tania.franca@embasa.ba.gov.br>

Fabrcio,

ANEXO C  
AUTORIZAÇÃO PARA USO DE DADOS



SECRETARIA DE INFRAESTRUTURA  
HÍDRICA E SANEAMENTO

Salvador, 19 de março de 2019

Atendendo a determinação da Diretoria Técnica, estamos autorizando e deliberando o estudante e funcionário, Ricardo Martins Carvalho, matrícula 12778-7, a ter acesso aos dados da base MAYA, Sistema Comercial e COPAE, referente a conclusão do curso de Análise e Desenvolvimento de Sistema.

Fabrício Mota Oliveira

Gerente da TRPA

## ANEXO D

### PRINTS DE TELA COM CARGA DAS DIMENSÕES

```
4200-ORGANIZACAO> Iteração sobre a consulta 1.1 realizada em 16ms
4200-ORGANIZACAO> Concluído em 16ms
4200-ORGANIZACAO> Iteração sobre a consulta 1 realizada em 1min 42s
4200-ORGANIZACAO> Concluído em 1min 42s
4200-ORGANIZACAO> Iteração sobre a consulta realizada em 1min 42s
4200-ORGANIZACAO> Concluído em 1min 42s
Macro 4200-ORGANIZACAO concluída em 1min 42s
Tempo total: 1min 42s
Pressione qualquer tecla para continuar. . .
```

Figura 19. Carga da Dimensão Unidade Organizacional

```
4200-ENDEREÇO> Consulta realizada em 93ms
4200-ENDEREÇO> Iteração sobre a consulta 1.1 realizada em 93ms
4200-ENDEREÇO> Concluído em 93ms
4200-ENDEREÇO> Iteração sobre a consulta 1 realizada em 5h 15min
4200-ENDEREÇO> Concluído em 5h 15min
4200-ENDEREÇO> Iteração sobre a consulta realizada em 5h 15min
4200-ENDEREÇO> Concluído em 5h 15min
Macro 4200-ENDEREÇO concluída em 5h 15min
Tempo total: 5h 15min
Pressione qualquer tecla para continuar. . .
```

Figura 20. Carga da Dimensão Endereço

```
4200-SERVIÇO> Iteração sobre a consulta 1 realizada em 5,24s
4200-SERVIÇO> Concluído em 5,26s
4200-SERVIÇO> Iteração sobre a consulta realizada em 5,62s
4200-SERVIÇO> Concluído em 5,62s
Macro 4200-SERVIÇO concluída em 5,62s
Tempo total: 5,62s
Pressione qualquer tecla para continuar. . .
```

Figura 21. Carga da Dimensão Serviço

```
4200-TEMPO> Iteração sobre a consulta 1 realizada em 20,1s
4200-TEMPO> Concluído em 20,1s
4200-TEMPO> Iteração sobre a consulta realizada em 20,4s
4200-TEMPO> Concluído em 20,4s
Macro 4200-TEMPO concluída em 20,4s
Tempo total: 20,4s
Pressione qualquer tecla para continuar. . .
```

Figura 22. Carga da Dimensão Tempo

```
4200-FATO> Repetição 6173: 1 registros lidos nesta consulta.
4200-FATO> Executando novamente...
4200-FATO> Iteração sobre a consulta 1.1 realizada em 15ms
4200-FATO> Consultando dados (db1-TCC.conf)...
4200-FATO> Retornado em 0ns
4200-FATO> Consultando dados (db1-TCC.conf)...
4200-FATO> Retornado em 16ms
4200-FATO> Consulta realizada em 16ms
4200-FATO> Iteração sobre a consulta 1.1 realizada em 16ms
4200-FATO> Concluído em 16ms
4200-FATO> Iteração sobre a consulta 1 realizada em 3min
4200-FATO> Concluído em 3min
4200-FATO> Iteração sobre a consulta realizada em 3min 1s
4200-FATO> Concluído em 3min 1s
Macro 4200-FATO concluída em 3min 1s
Tempo total: 3min 1s
Pressione qualquer tecla para continuar. . .
```

Figura 23. Carga da Tabela Fato.

### AGRADECIMENTOS

Agradeço a Deus acima de tudo. Aos meus pais, Agripino e Rosália (*In memoriam*), pelo amor, carinho, paciência e

ensinamentos que permearam meus caminhos para eu chegar até onde estou hoje. A minha esposa Michele, e nossa pequena Elis, minha família, amores de minha vida, e por quem eu enfrento tudo. A dona Iraci (*In memoriam*), uma segunda mãe para mim. Aos meus irmãos Agripino e Thaony por todo apoio e carinho. As minhas irmãs Adriana e Uilma, pelo cuidado, afeto e suporte. Em especial a meu irmão Samuel, amigo e grande incentivador, que sempre confiou em mim e me ensinou a nunca desistir. As minhas cunhadas Ivana e Débora pelo apoio, confiança. Aos colegas Fabricio Mota, Geusa Sales, Bruno Andrade, Fabio Souza, Catarina Martins, Ival Ferreira, que me deram contribuições preciosas no momento em que precisei. Ao professor Jowaner Araujo, pela ajuda constante e orientação nesse trabalho, além de contribuição fundamental na minha formação. A Paulo Ricardo, pela amizade e parceira nas atividades, você vai longe meu velho. Aos demais professores, por todo aprendizado proporcionado e, sobretudo, pelo exemplo a ser seguido. Enfim, todos que de alguma forma contribuíram nesse processo de formação acadêmica.

### REFERÊNCIAS

- [1] A. S. Piechnicki, J. L. Kovaleski, M. V. d. Souza, F. Piechnicki, and L. R. Baran, "Utilização da metodologia de análise e solução de problemas na redução das perdas de água: um estudo de caso na sanepar," *Revista de Engenharia e Tecnologia*, vol. 3, no. 2, pp. Páginas–90, 2011.
- [2] M. T. Tsutiya, *Abastecimento de água*. Departamento de Engenharia Hidráulica e Sanitária da Escola Politécnica da Universidade de São Paulo, 2004.
- [3] SABESP, "Redução de perdas, visão geral, 2007." pp. 1–26, 2007. [Online]. Available: <http://bit.do/esorF>
- [4] M. Montenegro, S. Neto, M. Onoyama, M. Thome *et al.*, "Demand management and loss water control in the federal district-brasil," in *The World Water Congress*, 22. IWSA, 1999, pp. 1–5.
- [5] L. HELLER and V. DE PADUA, *Abastecimento de água para consumo humano*, ser. Coleção Ingenium. UFMG, 2006.
- [6] Embasa, "Empresa baiana de Águas e saneamento s.a." 2018. [Online]. Available: <http://http://www.embasa.ba.gov.br/>
- [7] R. G. da Mota, "Importância da setorização adequada paracombate às perdas reais de água de abastecimento público," Master's thesis, Escola Politécnica da Universidade de São Paulo, 2010.
- [8] M. do Carmo Calijuri and D. Cunha, *Engenharia ambiental: Conceitos, tecnologia e gestão*. Elsevier Editora Ltda., 2013.
- [9] COPASA, "Companhia de saneamento de minas gerais," 2018. [Online]. Available: <http://www.copasa.com.br/wps/portal/internet>
- [10] M. do Desenvolvimento Regional, "Ministério do desenvolvimento regional," 2019. [Online]. Available: <http://www.mi.gov.br/desenvolvimento-regional>
- [11] SNIS, "Sistema nacional de informações do saneamento," 2016. [Online]. Available: <http://www.snis.gov.br/>
- [12] IWA, "Internacional water association," 2016. [Online]. Available: <http://www.iwa-network.org/>
- [13] ABES, "Associação brasileira de engenharia sanitária e ambiental," 2016. [Online]. Available: <http://abes-dn.org.br/>
- [14] Brasil, "Lei nº 11.445, de 5 de janeiro de 2007. estabelece diretrizes nacionais para o saneamento básico," *Diário Oficial da União*, 2007. [Online]. Available: [http://www.planalto.gov.br/ccivil\\_-03/\\_ato20072010/2007/lei/111445.htm](http://www.planalto.gov.br/ccivil_-03/_ato20072010/2007/lei/111445.htm)
- [15] A. Criminisi, C. Fontanazza, G. Freni, and G. La Loggia, "Evaluation of the apparent losses caused by water meter under-registration in intermittent water supply," *Water Science and Technology*, vol. 60, no. 9, pp. 2373–2382, 2009.
- [16] M. Vargas, *O negócio da água: riscos e oportunidades das concessões de saneamento à iniciativa privada : estudos de caso no sudeste brasileiro*. FAPESP, 2005.

- [17] E. Gonçalves and P. A. Alvim, *Pesquisa e combate a vazamentos não visíveis*, 2007.
- [18] Y. Çengel and J. Cimbala, *Mecânica dos fluidos - 3.ed.* AMGH Editora, 2015.
- [19] F. M. Oliveira, “Dbs,” 2010.
- [20] A. Berson and S. Smith, *Data Warehousing, Data Mining, and OLAP*, ser. Data warehousing/Data management. McGraw-Hill, 1997.
- [21] R. Kimball and M. Ross, *Data Warehouse Toolkit: Técnicas para Construção de Data Warehouses Dimensionais.* Makron Books, 1998.
- [22] R. Kimball, M. Ross, B. Becker, J. Mundy, and W. Thornthwaite, *The Kimball Group Reader: Relentlessly Practical Tools for Data Warehousing and Business Intelligence Remastered Collection.* Wiley, 2016.
- [23] R. GOLDSCHMIDT and E. PASSOS, *Data mining: um guia Prático.* Elsevier Editora, 2005.
- [24] EMBASA, “Copae - sistema de controle operacional de Água e esgoto,” 2016.
- [25] R. DE CARVALHO COSTA, *SQL Guia Prático - 2a edição.* BRASPORT, 2005.
- [26] R. Mitchell, *Instant Web Scraping with Java.* Packt Publishing, 2013.
- [27] A. A. B. de Normas Técnicas, “Nbr 12218,” 2017. [Online]. Available: <http://www.abnt.org.br/>