
Serviço Auto Escalável para Visualização de Dados Através de Containeres

Filipe Viana Monteiro^{* 1}

Abstract

A competitividade do mercado e a busca por operações mais lucrativas são alguns dos motivos que vêm obrigando as empresas a realizar decisões de negócio mais assertivas e em espaços de tempo cada vez menores. Nesse contexto, a utilização de dados vem se mostrando de grande valia, impulsionando a adesão de soluções de *business intelligence*. Essas soluções viabilizam a tomada de decisões embasadas em dados gerados pela empresa, seus clientes e até mesmo entidades externas ao negócio. Com o intuito de adequar a essa realidade, o TRE investiu no desenvolvimento de um modelo de referência capaz de implementar soluções de BI, abrangendo as etapas de transformação, armazenamento e visualização de dados. Esse trabalho tem por objetivo propor um aperfeiçoamento nesse modelo para garantir maior escalabilidade e disponibilidade dos serviços, através da implantação de *clusters* auto escaláveis de containers. O estudo de caso desse trabalho foca na adequação do componente de visualização de dados do modelo de referência.

1. Introdução

A maturidade alcançada por tecnologias como *big data*, inteligência artificial, internet das coisas, computação em nuvem, dentre muitas outras, e a redução dos custos básicos de TI, resultante da disponibilização de infraestrutura como serviço e a disputa de mercado entre os grandes provedores de serviços de computação (Grandinetti, 2013), vem viabilizando e impulsionando um movimento de transformação digital em diversos ramos e organizações.

As organizações, sejam elas públicas ou privadas, de pequeno ou grande porte, em sua maioria já operam com o auxílio de um conjunto de soluções de *software* que armazenam

os dados de negócio, gerados tanto pela própria organização quanto por clientes e parceiros. Somam-se a esses dados, as informações publicamente disponíveis na internet, de modo que esse conjunto de dados esconde informações valiosíssimas que se bem utilizadas, podem significar a sucesso de um negócio.

Aliando-se, então, a disponibilidade dos dados, a redução no custo dos insumos de TI e maturidade atingida por novas tecnologias, temos a construção de um cenário que vem impulsionando as iniciativas de desenvolvimento e implantação de soluções de dados por meio da aplicação de técnicas como *business intelligence* (BI), que consiste na aquisição, tratamento de dados de negócio e armazenamento de dados de negócio, bem como o processo analítico sobre esses dados com o intuito de descoberta de padrões e tendências (Biere, 2003).

Neste contexto, uma das iniciativas do Tribunal Regional Eleitoral do Rio Grande do Norte (TRE-RN) para fomentar o desenvolvimento de solução de dados foi a implantação de uma turma de residência em tecnologia da informação em parceria com a Universidade Federal do Rio Grande do Norte (UFRN), onde uma das ênfases a serem trabalhadas no Tribunal era exatamente a de *business intelligence*.

Esse formato de residência em TI trata-se de uma iniciativa pioneira da UFRN para, em parceria com órgão externos, fomentar a formação de profissionais qualificados em tecnologia e técnicas invadoras no âmbito da Tecnologia da Informação (TI). Neste programa, são selecionados alunos que cursarão um programa de pós graduação na Universidade e trabalharão alocados em projetos transformacionais dentro da instituição parceira, para ideação e desenvolvimento de soluções inovadoras por meio da aplicação de novas tecnologias.

Tratando-se especificamente da vertente de BI da turma de residência do TRE, e considerando que o TRE-RN não possuía modelo prévio, nem ferramentas voltadas para o desenvolvimento de produtos deste tipo, o objetivo principal da iniciativa foi a elaboração e implantação de um modelo de referência para desenvolvimento dos produtos composto por componentes de *software* livre.

Durante o curso desta residência, diversos *softwares*, técnicas e tecnologias foram testadas e avaliadas pelo corpo

¹Instituto Metr pole Digital, Universidade Federal do Rio Grande do Norte, Rio Grande do Norte, Brasil. Correspondence to: Filipe Viana Monteiro <filipevianam@gmail.com>.

técnico do Tribunal até que chegassemos em um modelo de referência, composto apenas de componentes gratuitos, que atendesse as necessidades e expectativas da equipe, resultando em uma solução capaz de suportar todo o processo de desenvolvimento de uma aplicação de *business intelligence*. Este modelo de referência tem seus componentes hospedados por meio de containers *Docker*, que já é um padrão de hospedagem de serviços do Orgão.

Com o intuito de garantir mais robustez e escalabilidade a esse modelo de referência, este trabalho descreve o processo de implementação de uma estratégia de escalabilidade automática e sob demanda de uma dos componentes desse modelo de referência desenvolvido durante a residência de TI do Tribunal Regional Eleitoral do Rio Grande do Norte.

A sessão dois deste artigo descreve o modelo de referência desenvolvido, apresentando os componentes de *software* que a compõe, bem como as motivações para suas escolhas. Posteriormente apresentam-se as ferramentas que darão suporte a esse processo de auto escalabilidade do serviço de visualização e, por fim, a implantação e os resultados dessa nova arquitetura são apresentados.

2. Modelo de Referência

O modelo de referência idealizado e implementado no TRE-RN é composto por componentes de *software* capazes de realizar as atividades referentes ao processo de desenvolvimento de aplicações de *business intelligence*, sendo elas: ETL, DW e visualização. A etapa de ETL (*Extract, Transform and Load*), corresponde ao processo de extração, transformação e carga dos dados oriundos das diversas fontes de negócio em um repositório de dados dedicado ao processo de análise. Os dados resultantes da atividade de ETL devem encontrar-se integrados e modelados em um formato mais adequado a carga imposta pelo processo analítico. Esse repositório mencionado é exatamente a segunda etapa de um processo BI, denominado *Data Warehouse*, que em tradução literal significa armazém de dados, e consiste no armazenamento dos dados produzidos pelo processo anterior. De posse dos dados já tratados e persistidos em um *Data Warehouse*, inicia-se a etapa de visualização dos dados, que consiste no desenvolvimento de painéis visuais responsáveis por apresentar as informações necessárias para a tomada de decisão de negócio.

As atividades descritas acima, são executadas no modelo de referência pelos seguintes componentes de *software*: *Pen-taho Data Integration*, realizando as atividades de ETL, PostgreSQL, atuando como um DW, e o *Metabase* sendo o componente onde são desenvolvidos e apresentados os painéis gráficos. O modelo de referência é ilustrado na Figura 1

Conforme indicado anteriormente, este trabalho demonstra



Figure 1. Arquitetura de BI no TRE-RN.

a implantação de um processo de orquestração de múltiplas instâncias para garantir a escalabilidade sob demanda, inicialmente, aplicado na camada de visualização, ou seja, do *Metabase*. Para que possamos realizar o objetivo exposto, precisamos descrever melhor a arquitetura do componente em questão, bem como, as demais ferramentas que serão utilizadas.

2.1. Metabase

O *Metabase* é uma aplicação *web*, desenvolvida e mantida por uma comunidade aberta, que consiste em uma aplicação *backend* que contém uma API REST, bem como os códigos para comunicação com os bancos de dados e realizar o processamento dos resultados das consultas, e uma aplicação *frontend* de página única (*Single Page Application*) responsável pelas interfaces de usuário do sistema, conforme descreve o "guia de desenvolvedor" (*Metabase*). Vale ressaltar que ambos o *backend* quanto o *frontend* da aplicação são encapsulados em um único pacote de *software*, que na realidade do TRE é hospedado por um container Docker. Além dos sub componentes já explicitados, o *Metabase* depende um sistema de gerenciamento de banco de dados (SGBD) que irá armazenar e servir os metadados da própria aplicação. A configuração definida pelo corpo técnico do Tribunal foi a utilização do PostgreSQL para servir este propósito, em função da existência prévia deste SGBD na instituição e a disponibilidade de colaboradores com *expertise* na ferramenta. A Figura 2 ilustra a configuração descrita acima.

Referências

- Biery, M. *Business Intelligence for the Enterprise*. Pearson Education, 2003. ISBN 0131413031.
- Grandinetti, L. *Pervasive Cloud Computing Technologies: Future Outlooks and Interdisciplinary Perspective*. IGI Global, 2013. ISBN 9781466646841.
- Metabase. Guia do desenvolvedor. URL <https://www.metabase.com/docs/latest/developers-guide.html>.

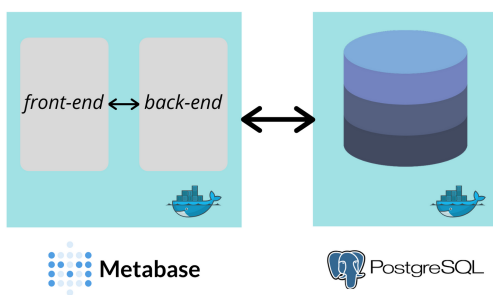


Figure 2. Arquitetura do Metabase.