

# Serasa Consumidor

Times do Serasa Consumidor

May 9, 2018

## 1 Desafio Serasa Consumidor

Antes de tudo, são duas páginas, mas não tem nada de difícil. Este problema tem o propósito de ser simples, rápido e objetivo. Nada de enrolação ou de ocupar muito do seu tempo. Queremos que este processo seja adequado a sua rotina e o instigue a concluí-lo.

Antes de explicar o desafio, vamos as poucas regras:

- Você tem total liberdade para usar as tecnologias que desejar. Claro, seria interessante se fosse utilizado tecnologias que estão dentro do nosso ferramental (Anexo I), mas isso não é obrigatório, afinal, tecnologias mudam, certo?!
- Queremos ao final do desafio, uma solução. E pra nós da Serasa Experian, uma solução tecnológica está bem acima de códigos ou compiladores. Soluções podem ser elaboradas em folhas de papel, desde que resolvam o problema.
- E sobre códigos: pra gente entender a qualidade do seu, precisamos vê-lo. Então, caso dê tempo, codifique tudo o que você conseguir. E se não der tempo, não se preocupe, foque em mostrar como você resolveria o problema com palavras.

## 2 O problema

Hoje, a **Serasa Experian**, como dito anteriormente, é o maior *Bureau* de crédito do Brasil. Aqui trabalhamos constantemente com grande volume e complexidade de dados. Sabendo disso, precisamos que você elabore uma solução que ofereça armazenamento, processamento e disponibilização desses dados, sempre considerando que tudo deve estar conforme as boas práticas de segurança em TI. Afinal, nosso principal ativo são dados sensíveis dos consumidores brasileiros.

### 2.1 Armazenamento

Vamos supor que existam três grandes bases de dados externas que organizam nossas informações. A primeira delas, que chamamos de *Base A*, é extremamente sensível e deve ser protegida com os maiores níveis de segurança, mas o acesso a esses dados não precisa ser tão performática. A segunda, é a *Base B* que também possui dados críticos, mas ao contrário da *Base A*, o acesso precisa ser um pouco mais rápido. Uma outra característica da *Base B* é que além de consultas ela é utilizada para extração de dados por meio de algoritmos de aprendizado de máquina. A última base, é a *Base C*, que não possui nenhum tipo de dado crítico, mas precisa de um acesso extremamente rápido.

### 2.2 Tráfego

Cada uma das bases existentes, são acessadas por sistemas em duas diferentes arquiteturas: micro-serviços e nano-serviços. Vale salientar que essas bases de dados são externas, portanto não é necessário dissertar sobre suas implementações, apenas suas consumações. Quanto aos payloads retornados por esses recursos, o candidato pode usar sua criatividade e definí-los, imaginando quais dados seriam importantes de serem retornados por sistemas como esses.

O primeiro sistema, acessa os seguintes dados da *Base A*:

- CPF

- Nome
- Endereço
- Lista de dívidas

O segundo, acessa a *Base B* que contém dados para cálculo do **Score de Crédito**. O **Score de Crédito** é um rating utilizado por instituições de crédito (bancos, imobiliárias, etc) quando precisam analisar o risco envolvido em uma operação de crédito a uma entidade.

- Idade
- Lista de bens (Imóveis, etc)
- Endereço
- Fonte de renda

O último serviço, acessa a *Base C* e tem como principal funcionalidade, rastrear eventos relacionados a um determinado CPF.

- Última consulta do CPF em um *Bureau* de crédito (Serasa e outros).
- Movimentação financeira nesse CPF.
- Dados relacionados a última compra com cartao de crédito vinculado ao CPF.

Como você resolveria esse problema? Divague sobre os seguintes tópicos e outros que ache adequado, sinta-se a vontade para desenhar, escrever, criar diagramas, vídeos, apresentação, ou qualquer outro meio que facilite o entendimento por parte dos avaliadores:

- Tecnologias adotadas
- Arquitetura utilizada
- Dados armazenados (já listados ou que você acrescentaria)

## 2.3 Disponibilização dos Dados

Agora que os dados desejados já foram consumidos, processados e armazenados, é necessário que eles sejam disponibilizados. Será necessário também desenvolver um meio pelo qual esses dados estarão disponíveis. É interessante imaginar os possíveis interessados em consumir esses dados para que uma única solução possa ser construída de modo a atender o máximo de situações possíveis.

## 3 Entrega da Solução

A solução deve ser entregue em um repositório público de sua preferência (**github**, **bitbucket**, etc). Deve conter um README explicando como utilizar qualquer código, caso exista. O repositório deve conter todos os arquivos utilizados para facilitar o entendimento da solução do problema de forma teórica. Não deve ser feita nenhuma referência à Serasa, SerasaExperian, SerasaConsumidor ou qualquer outra empresa do grupo Experian em variáveis, textos, nomes de banco, etc.

## 4 Anexo I

- Cloud AWS;
- Python, Node.js, Golang, PHP, ReactJs, Java, AngularJs, Groovy;
- Docker;
- PostgreSQL, ElasticSearch, Redis, MySQL, OracleDB, MongoDB;
- Terraform, Ansible, ShellScript, Jenkins;