



MBA ENGENHARIA DE DADOS

Integrantes:

Alessandro Bezerra Alves

Cristina Rodrigues Abrantes

Patrícia Claro

Acesso a arquitetura

Link: [Amazon Web Services \(AWS\) Sign-In \(signin.aws\)](https://signin.aws.amazon.com/signin?redirect_url=https%3A%2F%2Fgithub.com%2FHawaiiDataEngineers%2FSolutionSprint3)

Usuário: desafio-grupo-2

Senha: *G5W/M\$m

Informações da Arquitetura

A descrição da arquitetura, bem como scripts e as evidências de sua implementação encontram-se no github [HawaiiDataEngineers/SolutionSprint3 \(github.com\)](https://github.com/HawaiiDataEngineers/SolutionSprint3) resumizado como mostrado abaixo:

Solution Sprint 3

Repositório criado para a solution sprint 3 do curso de Engenharia de Dados.

Alunos:

- Alessandro Bezerra
- Cristina Abrantes
- Patricia Claro

Objetivo

Nosso objetivo é explicar e demonstrar a implementação da arquitetura de dados orientada a eventos proposta nesta solution sprint.

Índice

- [ENTREGAVEL 1 - Detalhes da arquitetura escolhida](#)
- [ENTREGAVEL 2 - Detalhes dos datasets escolhidos – URL, metadados, data, tamanho, entre outros](#)
- [ENTREGAVEL 3 - Implementação da arquitetura](#)
- [ENTREGAVEL 4 - Apresentação](#)



Backlog

Trata-se de ações que poderiam ser feitas para tornar ainda melhor a arquitetura proposta. São elas:

- Configuração de TTL para os buckets S3 Glacier e expurgo dos arquivos;
- Fazer a leitura e geração dos arquivos em memória e tempo de execução, não através arquivos em disco (exemplos arquivos json para o kinesis);
- Importância de um código python clean e instrumentado para redução do tempo de execução e facilidade de análise por outros profissionais;
- Otimização do código lambda que faz o handler do Kinesis para o envio de pacotes ao invés do envio linha a linha, reduzindo o número de iterações;