Solicitação do Usuário:

Eu como analista da Squad de Cartões

Solicito que seja criado um JOB DataBricks

Para leitura do arquivo CEXT Conforme regras abaixo:

========================

1. Arquivo será disponibilizado no Bucket S3 abaixo e deverá ser gravado no Bucket Bronze, como dado Bruto. Poderá ter mais de um arquivo na pasta, fazer o controle de arquivos lidos e arquivos que não atendam as regras dos itens 2 e 3
   1. s3:// Volumes/dev/tiintegracao/team/cartoes/cext/
   2. Arquivo modelo anexado a Estória (CEXT\_7562011\_20240125\_0002504.CCB)
2. Máscara para identificar arquivo CEXT, validar se o nome do arquivo obedece a regra abaixo:
   1. CEXT\_756\*.CCB
3. Validar estrutura do arquivo
   1. Primeira linha é cabeçalho e deve começar com = "CEXT0"
   2. Última linha é o trailer e deve começar com = "CEXT9"
   3. Criar uma tabela de log na camada bronze onde armazena-se a execução do JOB com sucesso ou com erros, com data e hora do início e fim de toda a execução, nessa tabela poderão ser salvos registros de outros JOBs. Deverá ter também os status de Em processamento, Sucesso ou Erro, caso erro, informar qual o erro.
4. Identificar se o lançamento é de uma conta do Ailos+
   1. Buscar número do cartão na linha na coluna 7 até a coluna 25 (19 posições)
   2. Conectar na base de dados do Ailos+
      1. Banco : Utilizar scripts DDL anexados pra criação das tabelas no banco de dados Oracle
   3. Executar a query:  
      SELECT 1  
        FROM CARTOES.TB\_CARTAO  
       WHERE NRCARTAO = &CARTAO;  -- Retorno do item 4.a
   4. Caso a query não retornar resultado, linha deverá ser descartada. Caso atenda a regra seguir par ao passo 5
5. Gravar dados do arquivo no Bucket Silver:
   1. Buscar ID do controle de arquivos  
      SELECT IDARQUIVO\_CONTROLE  
        FROM CARTOES.TB\_ARQUIVO\_CONTROLE  
       WHERE NRARQUIVO = 2;
   2. Caso tenha uma linha válida do item 4, efetuar o Insert na tabela CARTOES.TB\_ARQUIVO  
      INSERT INTO CARTOES.TB\_ARQUIVO  
        (IDARQUIVO,  
       IDARQUIVO\_CONTROLE  
        ,NMARQUIVO  
        ,DTARQUIVO  
        ,DHREGISTRO)

VALUES

  (id\_arquivo – Gerar um identificador único (raw(16))  
 ,id\_controle - Retorno do item 5.a  
  ,nome\_arquivo - Nome do arquivo processado  
  ,data\_arquivo - Buscar data do cabeçalho, posição 9 até a 16 (8 posições). Transformar de AAAAMMDD para DD/MM/AAAA e deverá ser uma data valida  
  ,data\_registro - data atual do sistema)

* 1. Insert na tabela CARTOES.TB\_ARQUIVO\_LINHA  
     INSERT INTO CARTOES.TB\_ARQUIVO\_LINHA  
       (IDARQUIVO\_LINHA   
      ,IDARQUIVO  
       ,NRLINHA  
       ,DSCONTEUDO  
       ,DTPROCESSO  
       ,CDSITUACAO)

VALUES

  (idarquivo\_linha - – Gerar um identificador único (raw(16))  
 ,id\_arquivo - Primary Key do insert do item 5.b  
  ,linha - Número da linha do arquivo  
  ,conteudo - Conteúdo da linha do arquivo  
  ,data - Data do sistema  
  ,1 - Fixo 1)

1. O JOB deverá ser disparado via chamada de uma API Databricks
2. Ao final do processo, espera-se ter os dados gravados nas camadas, Bronze, Silver e Gold conforme as regras descritas acima. A camada GOLD será um Banco de Dados Oracle conforme arquivo de estrutura enviada.
3. Fim do processo.
4. Deverão ser enviados os scripts Python, scripts da criação das tabelas nas camadas Bronze e Silver, bem como o JOB Criado e o Curl da chamada da API disparando o JOB

Desenho da arquitetura proposta:Interface gráfica do usuário, Aplicativo

Descrição gerada automaticamente