Anexo IV.2

Interface para Operação da Base de Controle

Data	Versão	Estado Atual	Autor
27/12/2024	1.0	Criação do documento	Signatárias
06/03/2025	2.0	Ajuste após OE Nº 6, Parecer Eletrônico Nº 4/2025	Signatárias
14/05/2025	3.0	Ajuste após OE Nº 15, Parecer Eletrônico Nº 12/2025	Signatárias

Sumário

1. CAPÍTULO I | V1

Referente às interfaces da Base de Controle da Versão 1 do SRO relacionadas ao registro do Seguro Garantia com obrigatoriedade definida pela SUSEP a partir de 2022.

1.1. OBJETIVOS

Proporcionar o entendimento do funcionamento de cada componente da Plataforma Integrada para atendimento a V1, como também o entendimento da própria Plataforma Integrada; e servir de referência para esclarecimento de dúvidas.

1.2. ESCOPO

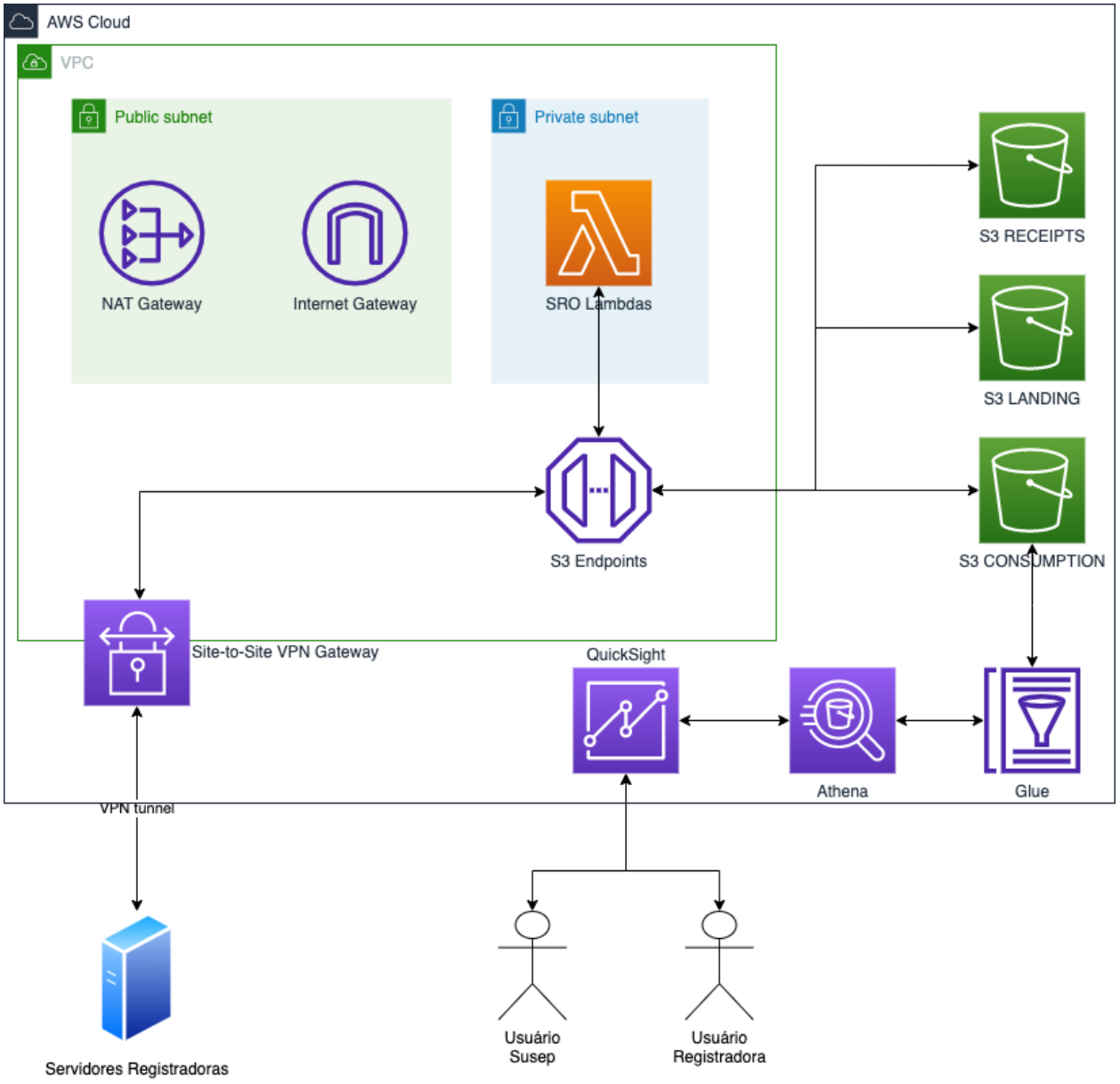
A documentação será dividida entre os seguintes componentes:

- Acesso ao ambiente de Produção;
- Modelos e leiautes firmados de arquivos para ingestão;
- Fluxo de push de arquivos pelas Signatárias em ambiente AWS;
- Estrutura de pastas em ambiente AWS de cada Signatária;
- Etapa de recebimento de arquivos;
- Etapa de validação de arquivos (Checklayout);
- Etapa de processamento de registros de arquivos conformes recebidos;
- Arquitetura de Multi-temperatura de Dados;
- Rotina de Unicidade;
- Dashboards de negócio e operacional; e
- Quadros estatísticos.

1.3. ACESSO AO AMBIENTE DE PRODUÇÃO

Uma vez a Plataforma Integrada em Produção, operada pelo Fornecedor de Tecnologia em atividade, o formato de acesso padrão para a ingestão dos dados de cada Signatária será através de VPN estabelecida entre a Signatária e a infraestrutura da Plataforma Integrada na AWS mantida pela Fornecedora atual, porém nesse primeiro momento algumas Signatárias efetivaram o acesso via liberação de IP's.

Abaixo, uma ilustração da rede da Plataforma Integrada:



1.4. MODELOS E LEIAUTES FIRMADOS DE ARQUIVOS PARA INGESTÃO

O leiaute é um documento utilizado para orientar o preenchimento dos dados pelos participantes e recebimento por parte das Signatárias. Através deste são aplicadas validações tanto as Signatárias quanto a Plataforma Integrada para ingestão dos dados .

O leiaute, na versão 1 do SRO cobre as seguintes movimentações exclusivamente para o Seguro Garantia: Apólice, Endosso, Liquidação, Contragarantia, Sinistro, Contrato Resseguro, Movimentação Prêmio Resseguro, Movimentação Recuperação Resseguro, Prestação Contas Resseguro e Exclusão, sendo que este último é o leiaute para o procedimento de exclusão de algum registro de algum outro leiaute. Ao todo são 10 leiaute.

O arquivo contendo os leiautes se chama LEIAUTES - SRO - SUSEP 20220107_V11.xlsx.

1.5. FLUXO DE PUSH DE ARQUIVOS PELAS SIGNATÁRIAS EM AMBIENTE AWS

A ingestão de dados é uma das etapas mais importantes da Plataforma Integrada. Nesta etapa as Signatárias enviam os arquivos com os registros do Participante para alimentarem a mesma, centralizando os dados para supervisão e consulta da SUSEP.

Na versão 1 do SRO, foi adotado a estratégia PUSH (envio de arquivos no formato Json) pelas Signatárias para suportar a ingestão dos dados no Data Lake da Plataforma Integrada.

A estratégia foi implantada através da utilização por cada Signatária de usuários IAM (Access key/Secret key), com acessos específicos às suas pastas, nos buckets do Data Lake da Plataforma Integrada para que possam enviar os arquivos, conforme leiautes definidos. Usando a credencial IAM, a forma de gerar e enviar os arquivos está a cargo de cada Signatária.

1.6. ESTRUTURA DE PASTAS EM AMBIENTE AWS DE CADA SIGNATÁRIA

A fim de termos uma organização com segurança em relação aos dados de entrada (Ingestão), enviados pelas Signatárias através dos arquivos em formato Json, ocorre a segmentação dos mesmos, dentro da estrutura do serviço S3 com acessos via credencial IAM.

A Plataforma Integrada está dividida em 3 buckets, sendo eles: bucket Landing para o recebimento dos arquivos enviados pelas Signatárias, bucket Receipts para a entrega dos recibos de recebimento, checklayout, processamento e unicidade dos arquivos enviados pelas Signatárias e bucket Consumption para acesso aos dados consolidados via Dashboards em QuickSight.

No bucket Landing, cada Signatária possui a sua pasta, onde através da sua credencial IAM, consegue acessar de forma exclusiva e com permissão somente de escrita, para efetuar o PUSH dos arquivos de ingestão de dados para a subpasta landing.

No bucket Receipts, cada Signatária possui a sua pasta, onde através da sua credencial IAM, consegue acessar de forma exclusiva e com permissão somente de leitura, para efetuar o resgate dos arquivos de recibos gerados pela ingestão de dados. Sendo que cada pasta de Signatária, possui 2 subpastas, receipts, onde aqui são entregues os arquivos de recibos derivados do processo de ingestão e unicidade e rejecteds, onde aqui são persistidos os arquivos originais usados na ingestão (para aqueles que foram invalidados no recebimento ou

no checklayout).

1.7. ETAPA DE RECEBIMENTO DE ARQUIVOS

Uma vez o arquivo de ingestão de dados existindo na subpasta landing, da pasta da Signatária no bucket Landing, aciona-se uma função lambda, onde esta efetua validações no nível do arquivo como, o nome do arquivo contém: cnpj, operação, data e hora, o cnpj do nome do arquivo é um cnpj válido e permitido, o tamanho do arquivo é menor que 64 MB, o arquivo é utf-8 ou compatível e o conteúdo do arquivo está em formato Json válido.

Caso haja alguma inconsistência, move-se o mesmo para a subpasta rejecteds da pasta da Signatária no bucket Receipts, gerando na subpasta receipts (da pasta da Signatária no bucket Receipts) um arquivo de recibo para o recebimento com inconsistência, sendo gerado notificação por e-mail a respectiva Signatária.

No caso de não haver nenhuma inconsistência, também é gerado um arquivo de recibo para o recebimento sem inconsistência e notifica-se por e-mail a respectiva Signatária.

A nomenclatura do recibo de recebimento é:

cnpj_Signatária_layout_data_hora_reciborecebido.json

1.8. ETAPA DE VALIDAÇÃO DE ARQUIVOS (CHECKLAYOUT)

No caso de tudo estar de acordo com a etapa de recebimento de arquivos, uma segunda validação ocorre, mas agora em relação aos dados contidos no arquivo, validando se cada tipo e formato de dado, em cada campo está condizente com a definição do leiaute se campos obrigatórios e condicionais estão preenchidos corretamente.

Caso haja alguma inconsistência, mesmo que em um único registro, move-se o arquivo inteiro para a subpasta rejecteds, da pasta da Signatária no bucket Receipts, gerando na subpasta receipts (da pasta da Signatária no bucket Receipts), um arquivo de recibo para o checklayout com inconsistência e notifica-se por e-mail a respectiva Signatária.

No caso de não haver nenhuma inconsistência, também é gerado um arquivo de recibo para o checklayout e sem inconsistência e notifica-se por e-mail a respectiva Signatária.

A nomenclatura do recibo de checklayout é:

- cnpj_Signatária_ layout_data_hora_recibochecklayout.json

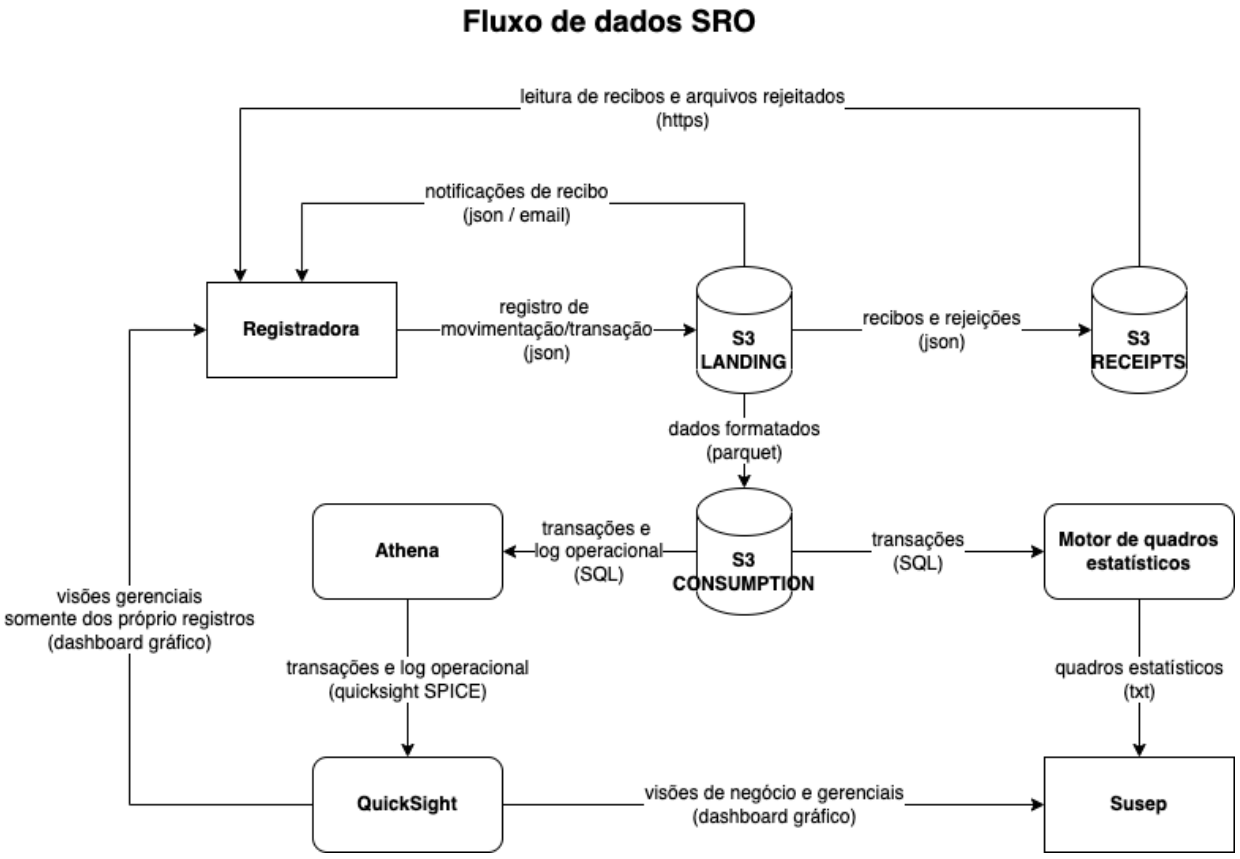
1.9. ETAPA DE PROCESSAMENTO DE REGISTROS DE ARQUIVOS CONFORMES RECEBIDOS

No caso de tudo estar de acordo com a etapa de validação de arquivos (CheckLayout), uma outra função lambda é acionada para converter o arquivo de ingestão de dados, do formato Json para o formato Parquet e alimentar o bucket Consumption, além de persistir o arquivo em partições de Signatária, Ano e Mês; visando facilitar e economizar no consumo dos dados, os Dashboards (QuickSight). Ao final deste processo a respectiva Signatária é notificada por e-mail.

A nomenclatura do recibo de processado é:

- cnpj_Signatária_layout_data_hora_reciboprocessado.json

Fig 1 - Processo do processamento de dados da V1



1.10. ARQUITETURA DE MULTI-TEMPERATURA DE DADOS

O processo de Data Aging no mundo clássico de banco de dados relacional, serve para a remoção de dados antigos do armazenamento. Porém este conceito em armazenamento do tipo Object Storage em sistemas modernos de dados em cloud tem a função de trazer mais

economia para o uso do armazenamento, podendo movimentar dados de um determinado período com menor acesso, para um local de armazenamento com custo reduzido, porém com acesso ao dado não tão instantâneo.

A AWS oferece as classes de armazenamento para o serviço S3, onde a classe de maior custo é a Standard, sendo esta a classe que oferece o maior nível de SLA de acesso 99,99%. Podemos configurar, baseado em períodos de data uma movimentação automática para os objetos (arquivos) de um determinado bucket com menor frequência de acesso para uma classe de armazenamento com um custo menor, porém com um SLA de resposta de acesso menor também.

O processo de multi-temperatura como está sendo chamado o Data Aging na Plataformra Interada SRO foi habilitado para os buckets Landing, Receipts e Consumption, sendo que em Landing e Receipts a movimentação não será para outra localidade e sim haverá o expurgo do arquivo da Plataforma Integrada depois de 30 dias da data em que o mesmo foi criado. A automação de expurgos de arquivos não foi habilitada no período de operação assistida, a pedido das Signatárias.

Já no bucket Consumption teremos a seguinte regra de movimentação: Para os arquivos com data de criação maior do que 365 dias, os mesmos serão movidos da classe Standard para a classe Standard-IA. Neste caso haverá somente redução de custo e não de tempo de acesso, pois ambas as classes de armazenamento possuem o mesmo SLA de resposta 99,99%.

1.11. ROTINA DE UNICIDADE

A rotina de Unicidade é composta por 2 processos, a não unicidade exata e a possível não unicidade e ambos são aplicados somente para o leiaute apólice.

O processo de não unicidade exata compreende a identificação de apólices em duplicidades onde as seguintes informações são iguais: Código da apólice, Código da seguradora, Código da filial, Tipo da emissão e número SUSEP da apólice. Esse contexto pode ser caracterizado pelo envio equivocado de mais de uma vez da mesma apólice pela Signatária, resultado possivelmente um erro operacional.

O processo de possível não unicidade já pode ser causado por uma situação de fraude, onde as seguintes informações são iguais: Código da seguradora, Tipo da emissão, Data de início da vigência da apólice, Data de término de vigência da apólice, Documento do segurado da apólice, Tipo do documento do segurado da apólice, Grupo e ramo da cobertura do objeto

segurado, Código da cobertura do objeto segurado e Valor total do prêmio em Real.

Ambos os processos são executados diariamente e sempre às 02h00 da manhã (Horário de Brasília) de todos os dias. Essas execuções, caso identifiquem não unicidade exata e/ou possível não unicidade, geram recibos cada uma e estes ficam disponíveis para resgate na subpasta receipts da pasta da Signatária no bucket Receipts.

Na situação de não unicidade (exata ou possível) inter-signatárias, ambas são notificadas e são gerados recibos para todas as Signatárias envolvidas, contudo, não há informação de quais as Signatárias coincidentes, de modo a assegurar o sigilo de informação entre as mesmas.

A nomenclatura dos recibos de não unicidade exata e possível não unicidade são:

- CNPJ_Signatária_AAAAMMDD_recibonaounicidade.json
- CNPJ_Signatária_AAAAMMDD_recibopossivelnaounicidade.json

1.12. QUADROS ESTATÍSTICOS

Foi implementada a rotina mensal de geração dos Quadros Estatísticos para os quadros 376, 377 e 378 (modelos SUSEP) com base nos dados contidos na própria Plataforma Integrada.

Os quadros são gerados mensalmente a partir do 11° dia do mês subsequente com dados da competência do mês anterior e disponibilizados em ambiente próprio.

Os Quadros Estatísticos são de uso exclusivo da SUSEP, gerados em formato texto, posicional com a seguinte nomenclatura de nome de arquivo AAA-MM-DD.txt, sendo que o DD é referente ao último dia do mês de competência e a SUSEP resgata-os na pasta quadros_estatisticos e subpastas q376, q377 e q378 do bucket Consumption.

1.13. OPERACIONAL

1.13.1 Interfaces

· QuickSight

Ferramenta provisionada dentro da Conta AWS para produção e visualização de dashboards;

· S3 Landing

Bucket do S3 (serviço AWS) destinado à recepção de arquivos do SRO;

· S3 Recibos

Bucket do S3 (serviço AWS) destinado à leitura de recibos e arquivos rejeitados do SRO;

- S3 Consumo

Bucket do S3 (serviço AWS) destinado ao armazenamento dos arquivos parquet do SRO, domínios e quadros estatísticos;

- Alarmes Operacionais

Alarmes do CloudWatch baseados em métricas operacionais das Lambdas que realizam as funções principais do SRO;

- Console do Athena

Console SQL do serviço Athena da AWS que permite a interação manual via queries SQL com o datalake do SRO;

1.13.2. Atores

- Entidade Signatária

Entidade responsável pela geração dos registros de movimentos associados ao mercado de seguros brasileiro;

- Usuário Signatária

Pessoa física, pertencente à Entidade Signatária;

- Entidade SUSEP

Entidade responsável pela absorção das informações geradas a partir do SRO;

- Usuário SUSEP

Pessoa física, pertencente à Entidade SUSEP;

- Fornecedora de Tecnologia

Entidade de administração técnica, responsável pela administração da conta AWS Susep na qual a versão produtiva do SRO se encontra;

1.13.1. Relação de Ações

A tabela abaixo traz uma relação das ações entre as entidades e as interfaces.

Ator	Interface	Ação
Entidade Signatária	S3 Landing	Upload de arquivo de registro
Entidade Signatária	S3 Recibos	Leitura de arquivos de recibo

		de movimentação
Entidade Signatária	S3 Recibos	Leitura de arquivos rejeitados no processo de ingestão
Usuário Signatária	QuickSight	Visualização de dashboards de negócio associados às movimentações da própria Signatária
Entidade SUSEP	S3 Consumo	Leitura dos arquivos referentes aos quadros estatísticos
Usuário SUSEP	QuickSight	Visualização de dashboards de negócio associados ao SRO como um todo
Usuário SUSEP	QuickSight	Visualização de dashboards gerenciais associados ao SRO como um todo
Fornecedora de Tecnologia	QuickSight	Manutenções evolutivas e correções
Fornecedora de Tecnologia	Conta AWS	Ajustes emergenciais
Fornecedora de Tecnologia	Console do Athena	Avaliação de problemas
Fornecedora de Tecnologia	Conta AWS	Manutenções evolutivas e correções
Fornecedora de Tecnologia	Conta AWS	Gestão de acessos das demais entidades
Fornecedora de Tecnologia	Alarmes Operacionais	Acompanhamento de problemas relacionados à infraestrutura dos motores de dados

Ajustes Emergenciais na Conta AWS pela Fornecedora

Contextos para a necessidade da ação:

- Presença de arquivo parquet inválido no DataLake, erro identificado através da constatação de congelamento da atualização dos datasets do QuickSight
- Fornecedora de Tecnologia reporta algum problema de segurança que afeta alguma interface da aplicação

Premissas para a execução da Ação:

- Mapeamento do plano de ajuste
- Fornecedora de Tecnologia realiza liberação temporária com permissões específicas

Descrição do processo:

Um usuário pré-determinado da Fornecedora de Tecnologia deve acessar a conta e realizar as ações conforme informado, acordado e aprovado pelas Entidades SUSEP.

Avaliação de Problemas no Console do Athena

Contextos para a necessidade da ação:

- Constatação de congelamento da atualização dos datasets do QuickSight
- Falha ou timeout na execução de alguma das Lambdas de msck repair

Premissas para a execução da Ação:

- Fornecedora de Tecnologia realiza liberação temporária com permissões específicas.

Descrição do processo:

Um usuário pré-determinado da Fornecedora de Tecnologia deve acessar o console e realizar as queries necessárias para a análise do problema reportado, conforme informado, acordado e aprovado pelas Entidade SUSEP. As consultas realizadas nesse contexto têm por objetivo apenas analisar o contexto funcional do DataLake e não realizar a extração de dados.

Gestão de Acessos das demais entidades na conta AWS

Os contextos para a necessidade desta ação são definidos pela Fornecedora

de Tecnologia. Premissas para a execução da Ação:

- Qualquer alteração de acesso não pode interferir na capacidade da aplicação de acessar os serviços da AWS às quais ela atualmente tem permissão via IAM Policy

Descrição do processo:

Com relação à aplicação SRO, a Fornecedora de Tecnologia deve informar, acordar e obter autorização da Entidade SUSEP para alterar a estrutura dos modelos de permissionamento dos Buckets S3 ou alterar os ids de VPC, Subnet ou Security Group, dado que a alteração dos mesmos pode impactar na funcionalidade do SRO.

Demais alterações como provisionamento de novos ranges de IPs nas políticas dos Buckets S3, gestão do IAM (desde que não envolva a deleção das políticas das Lambdas e do processo de DevOps), e provisionamento de interfaces de rede para as Signatárias, podem ser realizadas a critério das próprias e da Fornecedora de Tecnologia.

Acompanhamento de Problemas de Infraestrutura dos motores de dados nos alarmes operacionais

Contextos para a necessidade da ação:

- Essa ação deve ser executada em caráter contínuo.

Premissas para a execução da Ação:

- Os alarmes operacionais devem estar disponíveis na conta produtiva para utilização pelo time operacional da Fornecedora de Tecnologia.

Descrição do processo:

Serão criados alarmes referentes a características críticas de cada processo do SRO.

Abaixo segue uma relação destes Alarmes e sua criticidade em caso de ativação:

Alarme	Criticidade	Descrição
Erros do Processo de Ingestão	Alta	Este alarme indica falha no processo de ingestão, o que acarreta na não-ingestão de registros e na possível falha de informação do mesmo à Signatária que o gerou. A Fornecedora de Tecnologia deve ser informada imediatamente para avaliação e ajuste.

Erros do Processo de Atualização de Índices do Consumo	Média	Este alarme indica falha no processo de atualização de índices do Consumo, o que acarreta não-atualização de visualização no QuickSight. O processo pode ser realizado manualmente caso necessário.
Tempo de execução do Processo de Índices do Consumo	Baixa	Este alarme é um indicador de que a volumetria do projeto passou de um determinado ponto e pode acarretar num replanejamento de provisionamento computacional para a próxima manutenção evolutiva.
Erros do Processo de Atualização de Índices do Operacional	Média	Este alarme indica falha no processo de atualização de índices do Operacional, o que acarreta não-atualização de visualização no QuickSight. O processo pode ser realizado manualmente caso necessário.
Tempo de execução do Processo de Índices do Operacional	Baixa	Este alarme é um indicador de que a volumetria do projeto passou de um determinado ponto e pode acarretar num replanejamento de provisionamento computacional para a próxima manutenção evolutiva.
Erros do Processo de Mapeamento de Custo	Baixa	Este alarme é um indicador da falha da atualização de custo que aparece no QuickSight, o custo pode ser obtido diretamente na AWS caso necessário.
Erros do Processo de Geração de Quadros Estatísticos	Alta	Este alarme indica falha na geração de quadros estatísticos, a Fornecedora de Tecnologia deve ser informada imediatamente para avaliação e ajuste.
Tempo de execução do Processo de Geração de Quadros Estatísticos	Baixa	Este alarme é um indicador de que a volumetria do projeto passou de um determinado ponto e pode acarretar num replanejamento de provisionamento computacional para a próxima manutenção evolutiva.
Erros do Processo de Validação de Unicidade	Alta	Este alarme indica falha na validação de unicidade, a F deve ser informada imediatamente para avaliação e ajuste.

Tempo de execução do Processo de Validação de Unicidade	Baixa	Este alarme é um indicador de que a volumetria do projeto passou de um determinado ponto e pode acarretar num replanejamento de provisionamento computacional para a próxima manutenção evolutiva.
Erros do Processo de Ingestão de Logs	Média	Este alarme é um indicador de que houveram falhas na ingestão de logs operacionais. Estes podem ser obtidos manualmente mas a falha dos mesmos gera inconsistências no histórico na Visão Operacional do QuickSight. A Fornecedora de Tecnologia deve ser informada assim que possível para avaliação e ajuste.
Tempo de execução do Processo de Ingestão de Logs	Baixa	Este alarme é um indicador de que a volumetria do projeto passou de um determinado ponto e pode acarretar num replanejamento de provisionamento computacional para a próxima manutenção evolutiva.

2. CAPÍTULO II | V2 E V3

Referente às interfaces da Base de Controle da Versão 2 e 3 do SRO relacionadas ao registro dos dados com obrigatoriedades definidas pela Susep entre 2023 e 2026.

2.1. OBJETIVOS

Proporcionar o entendimento do funcionamento de cada componente da Plataforma Integrada para atendimento a V2 e V3, como também o entendimento da própria Plataforma Integrada; e servir de referência para esclarecimento de dúvidas.

2.2. ESCOPO

A documentação será dividida entre os seguintes componentes:

- Acesso ao ambiente de Produção;
- Modelos e leiautes firmados de arquivos para ingestão;
- Estrutura de pastas em ambiente AWS de cada Signatária;
- Etapa de recebimento de arquivos;
- Etapa de validação de arquivos (checklayout);
- Etapa de processamento de registros de arquivos conformes recebidos;
- Rotina de Unicidade;

2.3. ACESSO AO AMBIENTE DE PRODUÇÃO

Uma vez que a Plataforma Integrada esteja em Produção, ela será operada pelo Fornecedor de Tecnologia em atividade, o formato acesso à conta da AWS será realizado exclusivamente através do AWS Identity and Access Management (IAM). Cada Signatária deverá utilizar credenciais IAM específicas, que serão fornecidas pela Fornecedora, para acessar os recursos necessários na Plataforma Integrada AWS. Além disso, será necessário utilizar configurações de cross-account, como IAM roles para permitir que o ambiente produtivo da Signatária se comunique com a conta onde a Plataforma Integrada está hospedada na AWS.

Cada Signatária poderá acessar recursos em produção buckets S3, funções Lambda, banco Redshift (V3) e Athena (V2).

2.4. MODELOS E LEIAUTES FIRMADOS DE ARQUIVOS PARA INGESTÃO

O leiaute é um documento utilizado para orientar o preenchimento dos dados pelos participantes e recebimento por parte das Signatárias. Através deste são aplicadas validações

tanto as Signatárias quanto a Plataforma Integrada para ingestão dos dados.

V2: Layout: Versão Resumida Interop - 2023.09.20.xlsx

V3: Layout: Leiaute SRO v 3-0-0-rc-3.xlsx

2.5. ESTRUTURA DE PASTAS EM AMBIENTE AWS DE CADA SIGNATÁRIA

Cada signatária vai conseguir visualizar somente a sua respectiva pasta. O acesso será realizado através dos links conforme o exemplo abaixo:

<SIGNATÁRIA>

DEV-TRANSIENT-<SIGNATÁRIA>DEV-RAW--<SIGNATÁRIA>DEV-REPORTS--<SIGNATÁRIA>

HML-TRANSIENT--<SIGNATÁRIA>HML-RAW--<SIGNATÁRIA>HML-REPORTS--<SIGNATÁRIA>

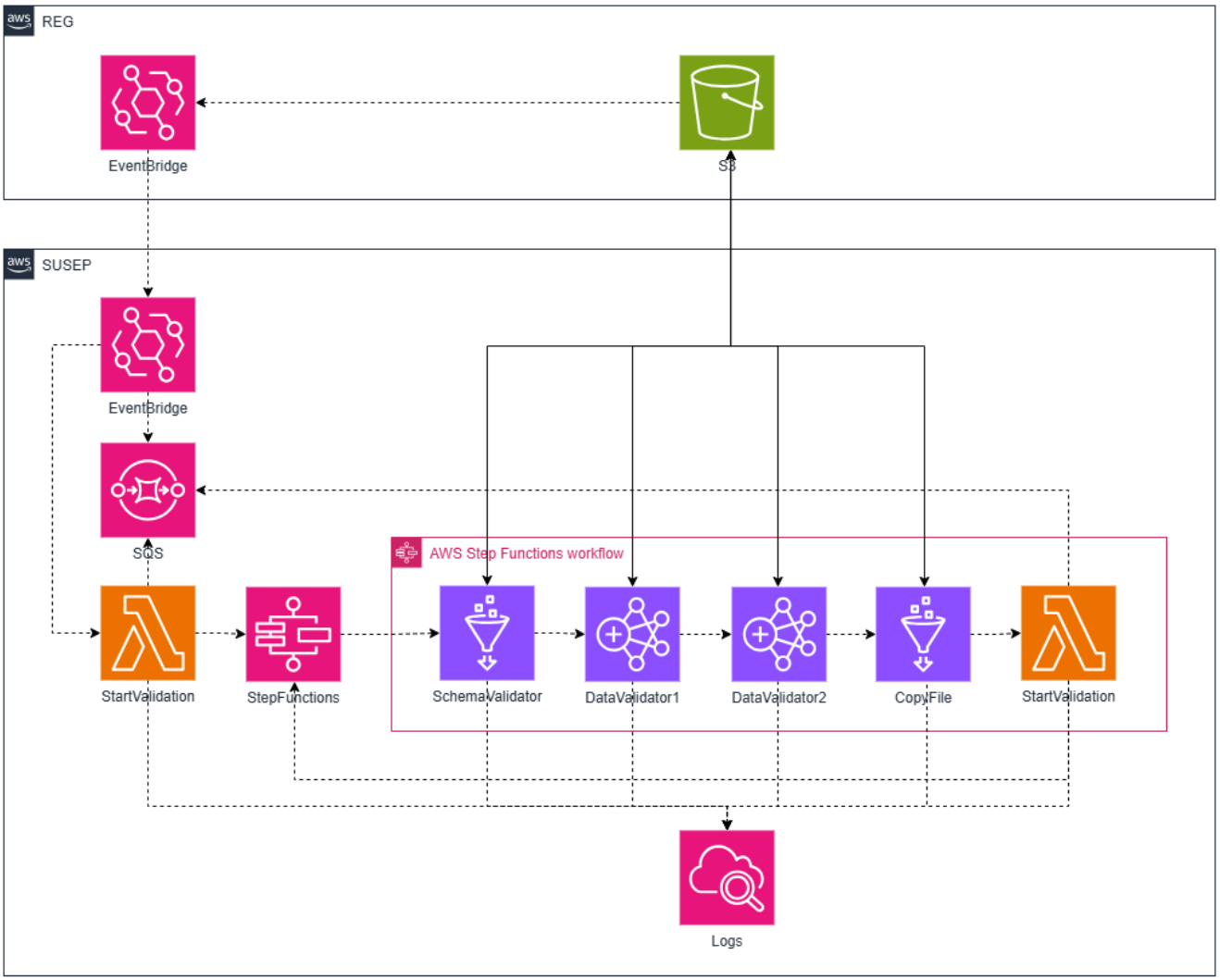
PRD-TRANSIENT-B3PRD-RAW--<SIGNATÁRIA>PRD-REPORTS--<SIGNATÁRIA>

2.6. ETAPA DE VALIDAÇÃO DE ARQUIVOS (CHECKLAYOUT);

2.6.1. Processo do processamento de dados da V2

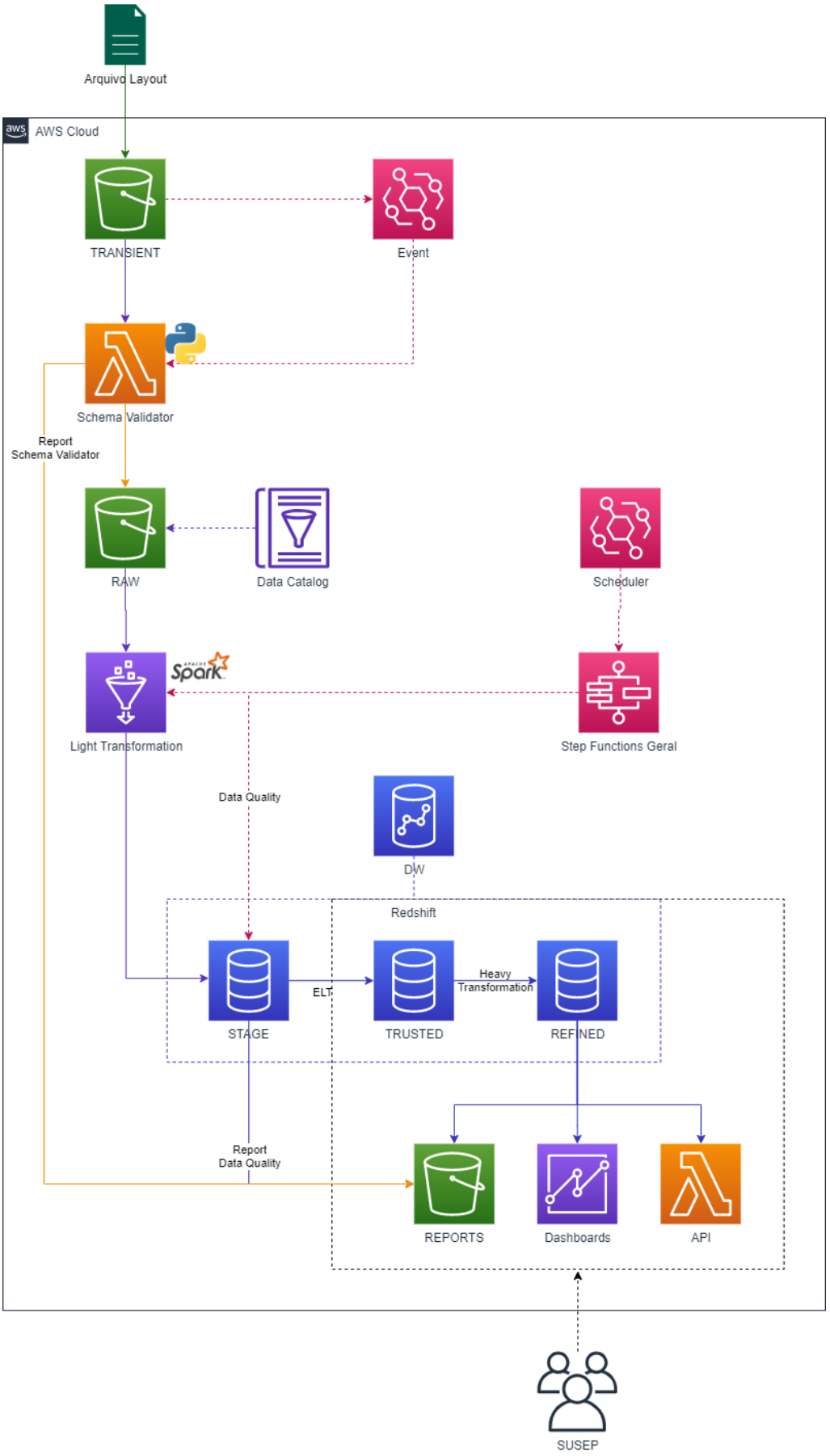
Os ambientes foram criados com base na arquitetura apresentada na Figura abaixo:

Fig 2 - Processo do processamento de dados da V2



2.6.2. Processo do processamento de dados da V3

Fig 3 - Processo do processamento de dados da V3



Atores

- Desenvolvedores das Signataria, SUSEP e Fornecedora de Tecnologia;
- Contas das Signatarias;
- Rede Cliente-Servidor

2.6.3. Parâmetros V2

Parâmetro	Valor
ANO	Ano de carga do arquivo
ARQUIVO	Nome do arquivo
BUCKET	Dia do processamento
CODIGO_SIGNATÁRIA	Nome do leiaute
DIA	Dia de carga do arquivo
MES	Mês de carga do arquivo
SIGNATÁRIA	Nome da Signatária
TABELA	• Ccg • Compl_auto • Documento • Endosso • Movimento_premio • Movimento_sinistro • Sinistro
UUID	ID aleatorio de execução

2.6.4. Parâmetros V3

Parâmetro	Valor
ANO	Ano do processamento
ARQUIVO	Nome do arquivo
DIA	Dia do processamento
LAYOUT	Nome do leiaute
MÊS	Mês do processamento
RAW	Nome do bucket da camada RAW
SIGNATÁRIA	Nome da Signatária
REPORTS	Nome do bucket de relatórios
TRANSIENT	Nome do bucket da camada TRANSIENT

2.7. ETAPA DE PROCESSAMENTO DE REGISTROS DE ARQUIVOS CONFORMES RECEBIDOS V2

2.7.1. Envio de Arquivos

Para submeter arquivos para validação eles deverão ser enviados para o diretório data_quality do Bucket, seguindo o padrão de diretório de cada leiaute. Exemplo:

s3://BUCKET/data_quality/TABELA/layout=LAYOUT/codigo_Signatária=CODIGO_SIGNATÁRIA/ano=ANO/mes=MES/dia=DIA/ARQUIVO

O evento de criação do arquivo será enviado e entrará em uma fila de processamento exclusiva para cada Signatária.

2.7.2. Ingestão e validação dos dados

Os dados são ingeridos por meio de eventos (Event Bridge) conciliados ao Função Lambda responsável por iniciar o processo de validação onde irá consumir os eventos do Amazon SQS e iniciar o orquestrador de data quality que é a Step Function.

2.7.3. Processamento

O arquivo será validado de acordo com as regras definidas pela SUSEP. Cada validação criará uma execução na StepFunction com o seguinte ID: SIGNATÁRIA-TABELA-ARQUIVO-UUID. A execução também pode ser acompanhada pelos logs no Cloudwatch, de acordo com o padrão /data_quality/SIGNATÁRIA/TABELA-ARQUIVO

1 – Os arquivos parquet são enviados no bucket Amazon S3 das Signatárias no qual a plataforma tem acesso através do Access Point.

1.1 - Arquivos enviados na pasta com nome de leiaute são inseridos diretamente no Redshift.

2.1 - Arquivos enviados na pasta com nome **/data_quality** são processados pelo data quality.

2 – O data quality é composto por 3 etapas de validação realizadas pelo EMR Serveless

2.1 - Schema Validator (Glue Job): Validação de estrutura do arquivo.

2.2 - Data Validator 1 (EMR Serveless): Valida obrigatoriedade, tamanho, formato e condições.

2.3 - Data Validator 2 (EMR Serveless): Valida regras entre blocos ou tabelas (Dependências)

3 – Após validação dos arquivos os Reports são gerados e armazenados no Bucket S3 na pasta **/report.**

2.7.4. Resultado

Durante o processamento poderão ser gerados até 3 relatórios de validação, que serão escritos no seguinte diretório:

- s3://BUCKET/report/ANO/MES/DIA/TABELA/ARQUIVO/schema_report.json
- s3://BUCKET/report/ANO/MES/DIA/TABELA/ARQUIVO/data_report1/
- s3://BUCKET/report/ANO/MES/DIA/TABELA/ARQUIVO/data_report2/

No final, caso o arquivo esteja correto, o mesmo será copiado para o diretório da tabela:

s3://BUCKET/TABELA/layout=LAYOUT/codigo_Signatária=CODIGO_SIGNATÁRIA/ano=ANO/mes=MES/dia=DIA/ARQUIVO

2.8. ETAPA DE PROCESSAMENTO DE REGISTROS DE ARQUIVOS CONFORMES RECEBIDOS V3

2.8.1. Envio de arquivos e validação de estrutura

Os arquivos deverão ser enviados para a camada TRANSIENT, seguindo o padrão de diretório de cada leiaute. Exemplo:

s3://TRANSIENT/SIGNATÁRIA/LAYOUT/ARQUIVO

A estrutura do arquivo será validada imediatamente e o relatório será gerado no caminho abaixo, na seguinte estrutura:

Validação de Estrutura:

s3://REPORTS/SIGNATÁRIA/data_quality/schema_report/LAYOUT/ANO/MÊS/DIA/ARQUIVO

Caso a estrutura do arquivo esteja correta, o mesmo será movido para a camada RAW:

s3://RAW/SIGNATÁRIA/LAYOUT/year=ANO/month=MÊS/day=DIA/ARQUIVO

2.8.2. Ingestão e Validação dos Dados

Uma vez por dia, de acordo com o agendamento definido, os arquivos do dia anterior serão processados e inseridos na camada STAGE do Redshift, onde serão validados. Cada registro será validado de acordo com as regras definidas na planilha de leiaute disponibilizada pela SUSEP. No final serão gerados dois relatórios de acordo com a definição em Relatórios e os registros corretos serão armazenados na camada TRUSTED do Redshift.

2.8.3. Processamento

1 – Os arquivos parquet são enviados para o bucket **Transient** na plataforma.

2 – Após arquivos inseridos através de eventos (EventBridge) é iniciado o processo de validação de estrutura dos arquivos, realizado por uma função AWS Lambda.

3 – Realizado a validação de estrutura o arquivo é inserido na **RAW** e disponíveis para consulta no AWS Athena

4 – Arquivo na RAW é iniciado o processo de ETL através do Amazon Glue Job e enviando o registro transformado para a **Stage** no Redshift

5 – Arquivo na Stage no Redshift é iniciado o processo de Data Quality

5.1 - A validação dos registros acontece no próprio Redshift validando as regras para cada campo (tamanho, formato, cardinalidade, condição, etc)

5.2 - Após validação é gerado arquivos de Reports no bucket de Report

6 – Finalizado a validação é iniciado a etapa de ELT onde os arquivos são transformados e armazenados na cama **TRUSTED** no Redshift

2.8.4. Resultado

Durante o processamento poderão ser gerados até 3 relatórios de validação, que serão escritos no seguinte diretório:

- s3://BUCKET/report/ANO/MES/DIA/TABELA/ARQUIVO/schema_report.json
- s3://BUCKET/report/ANO/MES/DIA/TABELA/ARQUIVO/data_report1/
- s3://BUCKET/report/ANO/MES/DIA/TABELA/ARQUIVO/data_report2/

No final, caso o arquivo esteja correto, o mesmo será copiado para o diretório da tabela:

s3://BUCKET/TABELA/layout=LAYOUT/codigo_Signatária=CODIGO_SIGNATÁRIA/ano=ANO/mes=MES/dia=DIA/ARQUIVO

2.9. PROCESSO DE UNICIDADE

O processo de não unicidade exata compreende a identificação de apólices em duplicidades onde as seguintes informações são iguais: Código da apólice, Código da seguradora, Código da filial, Tipo da emissão e número SUSEP da apólice. Esse contexto pode ser caracterizado pelo envio equivocado de mais de uma vez da mesma apólice pela Signatária, resultado possivelmente um erro operacional.

2.10. MODELOS E LEIAUTES FIRMADOS DE ARQUIVOS PARA INGESTÃO

O leiaute Padrão SRO é um documento utilizado para orientar o preenchimento dos dados pelos Participantes e recebimento por parte das Signatárias. Através deste são aplicadas validações tanto as Signatárias quanto a Plataforma Integrada para ingestão dos dados .

O leiaute , na versão vigente do SRO, cobre as seguintes movimentações : Documento, Documento Alteração, Sinistro Evento Gerador , Sinistro Alteração, FIE Provisão, Resgates e Portabilidades, CTT Assist, Alteração CTT, Alteração Saldo Devedor CTT ,Cosseguro Aceito, Alteração Cosseguro Aceito, Sinistro Coss Aceito, Alteração Sinistro Coss Aceito, Contrato Resseguro, Alt Contrato Resseguro, Prêmio Resseguro, Alt Prêmio Resseguro , Sinistro Resseguro, Alt Sinistro Resseguro, Bloqueio e Gravame , Transferência, Portabilidade, Capitalização e Exclusão, sendo que este último é o leiaute para o procedimento de exclusão

de algum registro de algum outro leiaute. Os dados de Capitalização e Resseguro serão trabalhados futuramente

Deve-se seguir o leiaute vigente na SUSEP, em acordo com o Termo de Adesão, com suas revisões periódicas.

3. ESTRUTURA DE ARMAZENAMENTO DE ARQUIVOS – PASTAS, RECEBIMENTO, PERSISTÊNCIA E EXCLUSÃO DE ARQUIVOS

A arquitetura possui diferente camadas de armazenamento ao logo das etapas de processamento:

Recebimento

A camada *TRANSIENT*, construída no S3, é responsável por receber os arquivos enviados e armazená-los temporariamente, por 7 dias, separados por Signatária e leiaute. Quando um arquivo for recebido, ele terá sua estrutura validada e, caso esteja correto, será movido para a camada RAW;

Armazenamento de Arquivos

A camada *RAW*, também construída no S3, é responsável por armazenar os arquivos validados temporariamente, por 60 dias, separados por Signatária, leiaute e data de envio;

Ingestão e Validação de Dados

Uma vez por dia, à meia-noite, será iniciado o processamento para ingestão dos dados do dia anterior contidos na camada *RAW*. Os dados de cada leiauteserão divididos de acordo com seus respectivos blocos e inseridos em tabelas na camada *STAGE*;

A camada *STAGE* está dentro do Redshift e armazena os dados temporariamente, desde a última execução, separados por Signatária, leiautee blocos;

Na camada *STAGE* será realizada a validação dos dados;

A validação de unicidade é uma das validações realizadas nessa etapa, que possui todas as validações que foram definidas na planilha de leiautepadrão SRO. Após cada execução serão gerados 2 relatórios de falha e de sucesso. O relatório de falha contém todos o registros que falharam em alguma regra, bem como o detalhamento da respectiva regra. O relatório de sucesso contém tdos os registros que passaram em todas as regras.

Persistência

Os dados válidos serão persistidos na camada *TRUSTED*, também no Redshift, e ficarão disponíveis para consulta;

Exclusão

A etapa seguinte é o processo de exclusão, responsável por verificar quais registros devem ser “excluídos” de acordo com os envios do leiautede exclusão.

Por último, os dados serão processados e salvos na camada *REFINED*, no Redshift, para uso em Dashboards, Relatórios e APIs.

Assinado por:
Eduardo Juliano
1055D70EFBBA4BB...

CIP S.A.

Assinado por:
Celso Lanzelotti
454081741CC4434...

MAPS Services S.A.

Assinado por:
Marcelo Rodrigues Costa
2140CFCE74DC479...

B3 Brasil, Bolsa, Balcão

DocuSigned by:
Gabriel Lorandos Germani
434CC0DA090746E...

Central de Serviços de Registro e Depósito aos Mercados Financeiro e de Capitais S.A