



- Considerando que os usuários dos aplicativos e website estão no Brasil, a arquitetura contempla a hospedagem dos serviços dentro da Region São Paulo. Desta forma, ganhamos em performance diminuindo a latência e o tempo de resposta das requisições;
- Para garantia de uma aplicação de alta disponibilidade e confiabilidade, foi projetado loadbalancer para multi Availability Zones, sendo uma dedicada para o ambiente de disaster recovery, além de autoscaling dos recursos, que também contribui para a otimização dos custos com a nuvem (OPEX), uma vez que o autoscaling mantém o sizing (horizontal e vertical) de acordo com os workloads que estão sendo consumidos, sem deixar recursos ociosos;
- O autoscaling é acionado de acordo com regras de alarmes parametrizados no Cloud Watch. Além disso, está previsto funções lambda ligadas ao Elasticsearch e eventualmente a um Kibana para coleta de métricas com o objetivo de monitorar as aplicações;
- Os microsserviços serão hospedados dentro de um cluster ECS e instâncias EC2 e as imagens serão registradas em repositório do ECR;
- Banco de dados DynamoDB para suporte de altas cargas de trabalho e requisições;
- O frontend será armazenado em bucket do S3 conectado ao CloudFront para distribuição do conteúdo. Pode-se configurar edge locations caso a empresa resolva expandir globalmente, dessa forma, usuários ao redor do globo terão uma experiência de acesso mais elevada;
- AWS IAM será utilizado para o gerenciamento das políticas, roles, e controles de acesso;
- AWS Secrets Manager configurado para gerenciar e proteger credenciais, tokens e security keys de APIs and banco de dados;