INSTITUTO DE TECNOLOGIA E LIDERANÇA – INTELI

Automation and Cost Management Project Based on TAGs in AWS

Public Report – Módulo 3

Authors:

Gustavo Monteiro

Advisor: Rodolfo Riyoei Goya

SÃO PAULO – SP

06/10/2025

1. Introduction

Building upon the formalized Tagging Standard established in Module 2, this third phase of the project transitions from governance definition to practical implementation. The primary objective of this module was to develop an automated system for collecting and exposing cost and resource data, thereby operationalizing the tagging strategy defined in partnership with Thomson Reuters. This involved creating a functional API, designing a resilient architecture, and validating its scalability. A key success factor in this phase was an agile approach, emphasizing proactive planning and early identification of potential impediments, which ensured that this technically demanding module was executed smoothly and effectively.

2 . Proactive Planning and Initial API Development

The module began with a crucial strategic planning session with the project advisor to map out the integration between the proposed API and the AWS platform. This foresight was instrumental, as it allowed the team to anticipate and mitigate potential blockers before implementation began. By addressing complex integration questions upfront, the development process proceeded with minimal friction.

The initial technical work focused on creating a foundational API using the FastAPI framework, chosen for its performance and auto-generating documentation features. Secure management of AWS credentials was implemented from the start using environment variables, preventing sensitive information from being hardcoded.

During this stage, a challenge arose when the AWS Academy lab account was temporarily blocked due to resource usage that exceeded the lab's established limits. This issue was promptly resolved in collaboration with the advisor, serving as a practical lesson in cloud environment constraints without causing significant project delays.

3. Collaboration and Iteration with the Business Partner

The project's architecture evolved significantly throughout the module to meet the demands of a real-world production environment.

3.1. From Proof-of-Concept to a Database-Backed API

The initial script was transformed into a robust two-component system. A Collector Component (populate_db.py) was developed to fetch data from AWS APIs periodically. This data was then stored in a local SQLite database (tcc_aws.db). The API Component (app.py)

was then refactored to serve data directly from this local database, resulting in faster response times, reduced dependency on live AWS API calls, and the creation of a historical data record.

3.2. Transition to a Cloud-Native Serverless Architecture

To eliminate the need for manual script execution and create a fully autonomous system, the architecture was further evolved into a serverless model leveraging core AWS services:

- **AWS Lambda**: A function was created to contain the data collection logic, running the code without the need to manage servers.
- **Amazon S3**: An S3 bucket was established as a centralized, durable data store for the JSON files (e.g., costs.json, instances.json) generated by the Lambda function.
- **Amazon EventBridge**: A rule was configured to automatically trigger the Lambda function once every 24 hours, ensuring the data in S3 remains consistently up-to-date without human intervention.

In this final architecture, the FastAPI application accesses the data directly from S3, providing a solution that is decoupled, resilient, and highly scalable.

4. Validation and Large-Scale Cost Analysis

To validate the system, a test environment was deployed using AWS CloudFormation. This simulation mimicked a corporate setting with multiple products and environments, and critically, included a "Non-Compliant Resource", an untagged EC2 instance, to successfully test the system's ability to identify governance failures.

Furthermore, a cost analysis was conducted for a large-scale corporate scenario (tracking 25,000 resources). The analysis projected a total monthly cost for the entire serverless data collection architecture of approximately $7.93 per month. This extremely low operational cost validates the financial viability and efficiency of the chosen serverless approach for enterprise-scale implementation.

5. Challenges and Next Steps

The main challenge encountered was the temporary suspension of the AWS lab account, which was successfully managed and served as a valuable learning experience regarding cloud platform limitations.

The agile methodology, focused on resolving difficult issues proactively, proved to be the key to this module's success. What was anticipated to be the most challenging phase was executed efficiently due to this strategic foresight.

With an automated and cost-effective data collection system now in place, the clear next step for the project is to focus on data visualization. The subsequent module will be dedicated to developing dashboards and reports. This will transform the collected data into actionable insights for stakeholders, allowing them to easily analyze costs, track resource distribution by tags, and identify non-compliant assets.

6. Professional Development and Practical Application of Knowledge

This module was a relevant learning experience, bridging the gap between theoretical concepts and tangible, professional application. The proactive planning sessions reinforced the immense value of addressing difficult questions early in the development lifecycle. This approach was the key to transforming a potentially problematic module into a successful and fluid execution.

Developing the API and evolving its architecture from a local script to a full-fledged serverless system provided hands-on experience with critical cloud services like AWS Lambda, S3, and EventBridge. This practical implementation solidified my understanding of how to build scalable, resilient, and remarkably cost-effective solutions in the cloud, offering a powerful strategic perspective that I can apply directly to my professional work.