UNIVERSITÀ DEGLI STUDI DI MILANO

FACOLTÀ DI SCIENZE E TECNOLOGIE

Corso di Laurea Magistrale in Informatica

SERVERLESS APPLICATION FOR MULTIPLE DATEBASES GOVERNANCE
ON AWS

Relatore:
Elena PAGANI
Correlatore:
Roy S. HALSTEAD

Tesi di Laurea di:
Michael DANIEL NAGUIB
Matricola: 923425

Anno Accademico 2020/2021

Data Entry Tool

The problem of centralizing the management of large-scale manufacturing data plagues many companies that have decided to embark on the Digital Transformation of their infrastructure. For this thesis project, I was hired by Deloitte Touche Tohmatsu, one of the Big Four in auditing and consulting, to work on the Cloud Transformation process of a major client company in the manufacturing sector. This transformation consisted of creating a cloud-based platform for the company's Data Management department, where users could seamlessly interact with databases distributed nationally or internationally.

The existing infrastructure that was to be replaced consisted of an on-premise data platform that was not designed for the amount of data being generated across all of the client's factories scattered throughout the territory. This solution was not flexible with respect to the various databases it had to manage, because for each new database, and for each new table associated with it, a new interface had to be implemented to manage the data. Furthermore, it did not place any control on the quality of the data entered in the various tables, did not provide any possibility to extract the data for use elsewhere, did not place any restrictions on what operations were available, nor to whom these tables were accessible. Continuing to do development for this platform had become unsustainable for the client, as each new page that had to handle a single table took a long time to develop, as well as being a very expensive process.

The Data Entry Tool project stands as an alternative to the data entry and manipulation platform. The new web application that was to be developed had to be a plug and play solution that could handle any connected database, it had to provide a system to control access to tables, along with the ability to group them as needed, it had to provide a level of validation for data types and impose any constraints on values, give the ability to import and export data in bulk, while also providing a logging system for each operation performed on the data. All of this should be available in a graphical interface that adapts to user permissions and table structures.

In my thesis, I introduce the concepts behind Cloud Transformation, such as different infrastructure models and service models. Emphasis is placed on microservices infrastructure and Serverless architecture, as they lie at the heart of the Data Entry Tool. Next, I describe the tools used within the project, from the Cloud Provider to the programming language chosen for development. Going forward, I present a deep dive of the implemented architecture, going from the Back-End and Front-End sub-architectures to the deployment system of the entire application. In addition, I explain the most important elements of the implementation I have done. Here, I explain how in the Back-End section I mainly dealt with managing multiple databases, collecting data to facilitate the dynamic creation of tables, and defining the schema used by the database that manages the application, while for the Front-End section I created a standard structure to manage static interfaces; moreover, I dealt with the Continuous Integration and Continuous Delivery of the application.

Towards the end, I show some implementation gimmicks and the impacts they had on application performance, along with comparisons between the implemented solution and the pre-

existing one to verify that the newly developed application has met the requirements. Finally, I present the final considerations, as well as some references to possible future developments.