



INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA
Department of Engenharia de Electrónica e Telecomunicações e de
Computadores

IoT System for pH Monitoring in Industrial Facilities

47185 : Miguel Agostinho da Silva Rocha (a47185@alunos.isel.pt)
47128 : Pedro Miguel Martins da Silva (a47128@alunos.isel.pt)

Report for the Curricular Unit of Project and Seminar
of Bachelor's Degree in Computer Science and Software Engineering

Supervisor: Doctor Rui Duarte

Abstract

In Germany there are several heavy restrictions on the discharge of waste resulting from manufacturing activities into nature, Therefore, our group has teamed up with the German company *Mommertz GmbH* which manufactures boiler filters to regulate the acidity of the boiler discharge liquid before it is released into the public sewage system. The issue at hand is that, over time, these filters experience deterioration over time due to usage, leading to a decline in their original filtering capacity, necessitating regular servicing. Currently, the company's strategy for dealing with this problem is to advise customers to periodically test the pH of the filter unit, and if the pH is lower than 6.5 at the point where the filtered material exits the filter, maintenance will be required for that unit. In this project, we implement automation for inspecting the pH filters. This solution is based on an embedded system and server-side computation. We choose which components to use to make each sub-system, and at the present moment we have a system that can read pH of the environment in the filter and can share that data to a central server which will process the data and send a signal to the manager of the device, via email, indicating that a value surpassed a threshold.

Table of Contents

1	Introduction	8
1.1	Motivation	8
1.2	Objectives	9
1.3	Document Organization	10
2	Background	11
2.1	IoT Systems	11
2.1.1	Sensors	11
2.1.2	Embedded Systems	13
2.1.3	Backend	16
2.1.4	Communication protocols	18
2.1.5	Frontend	19
2.2	Existing solutions	19
2.3	Summary	20
3	Proposed IoT System Architecture	21
3.1	System Overview	21
3.2	IoT Node	22
3.2.1	MCU	23
3.2.2	Sensor Modules	24
3.2.3	Sensor readings - Parallelized vs Serialized	25
3.2.4	Sensor Calibration	25
3.2.5	Ph Sensor	29
3.2.6	Assembly	33
3.2.7	Wi-Fi	34
3.2.8	MQTT	35
3.2.9	Storage	36
3.2.10	Time	36

3.3	Backend	36
3.3.1	Broker	37
3.3.2	Application Server	37
3.3.3	System Logic	38
3.3.4	Sensor Analysis	39
3.3.5	Spring Configuration	40
3.3.6	Authentication and Authorization	40
3.3.7	Controllers	41
3.3.8	Database	41
3.4	Website	43
3.4.1	Website Behaviour	44
3.4.2	Protected Resources	45
3.4.3	Error Handling	45
3.4.4	Services	46
3.5	System Tests	47
3.5.1	Backend	47
3.5.2	MCU Firmware	47
3.6	Security Measures	50
3.6.1	Password Storage	50
3.6.2	Website	51
3.7	Deploy	52
4	Conclusion	54
4.1	Future Work	54
4.1.1	Hardware Adaptation	54
4.1.2	Dynamic Configuration	54
4.1.3	pH Transformation with Temperature Compensation	55
4.1.4	Hermetic Enclosure	55
4.1.5	Scalability	55
4.1.6	Water Flow sensor	55
4.1.7	Summary	55

List of Figures

2.1	DHT11 Sensor Circuit	12
2.2	DHT11 Data Timing Diagram	12
2.3	Espressif ESP32-S2-SAOLA-1M (WROOM)	16
3.1	Main system architecture	21
3.2	IoT Node Module Architecture	22
3.3	Sliding window algorithm	26
3.4	Standard Deviation over Time	27
3.5	Standard Deviation over Time	28
3.6	Voltage divider circuit used to achieve lower output voltage	30
3.7	Graph illustrating the measured voltage with the MCU against the real voltage measured with an oscilloscope, of a liquid solution, pH equal to 8.	30
3.8	Graph illustrating effect of differing reference voltages on the ADC voltage curve.	31
3.9	Graph illustrating the measured ADC and voltage of a liquid solution, pH equal to 4.	32
3.10	Graph illustrating the measured humidity and temperature, from with the DHT11 sensor, at room temperature.	32
3.11	Graph illustrating the measured ADC and voltage when the sensor is immersed in water and taken out.	33
3.12	Current IoT Node assembly in the breadboard	35
3.13	Spring server overview	38
3.14	Postgres DB Entity Model	42
3.15	Time Series DB representation	43
3.16	Single Page Application components diagram	44
3.17	Authentication control in the SPA	46

3.18 Graph illustrating the ESP32-S2 energy consumption measured when the MCU wakes up to read from all sensors	48
3.19 Visual representation of the function of the salt [1]	51
3.20 Containers for each system service layer	52

List of Tables

2.1	Comparison Between different MCUs from espressif family, prices consulted from Mouser Online Store at 10/03/2023, URL: https://pt.mouser.com/	14
2.2	MSP430FR413x analysis, price consulted from Mouser Online Store at 10/03/2023, URL: https://pt.mouser.com/	14
3.1	Window Size vs. Time to Stabilize	28
3.2	Results for the different MCU consumption stages	48
3.3	Results for the different MCU profiles	49

1

Introduction

This document describes the implementation of the IoT System for pH Monitoring in Industrial Facilities project, developed as part of the final project in the Bachelor in Computer science and engineering degree[2].

1.1 Motivation

In Germany, there are strict laws for the discharge of water into the sewage. This is due to the protection of residual waters, and the sewage pipes, that conduct them. Any entity that acquired gas boilers, is obligated to neutralize the pH of water before discharging it into the public sewage system.

In the course of our final project, we collaborated with a German company called Mommertz GmbH[3]. The main aim of Mommertz's business model is to produce specialized filters designed to neutralize the pH of water resulting from the action of gas boilers, thus ensuring compliance with legal obligations.

The pH scale is a measure of the acidity or alkalinity of a solution. It is a logarithmic scale that indicates the concentration of hydrogen ions (H^+) in a solution. The pH scale ranges from 0 to 14, where a pH of 7 is considered neutral. Values below 7 indicate acidity, with lower numbers indicating higher acidity, while values above 7 indicate alkalinity or basicity, with higher numbers indicating higher alkalinity[4].

The filter comprises granules that serve to neutralize the pH of the water, including the water flowing from the boiler and passing through the filter. This system, like any industrial product, may suffer from several problems during its lifetime. As the water flows and undergoes neutralization, the granules gradually decompose, resulting in a reduction in the effectiveness of neutralization. Consequently, the pH of the water will not be fully neutralized.

Conversely, if an excessive amount of granules is used, the pH of the water may become excessively high. Furthermore, over time, water leaks may occur, and there is also a possibility of unexpected interruptions in water flow, necessitating maintenance. Currently, all the inspection for possible problems in the mechanism is done manually, which poses significant drawbacks as these types of systems are often situated in hard to reach locations.

Additionally, manual inspections result in needless maintenance and are prone

to human error. This heavy human dependency in the inspection for a possible malfunction in the behaviour of the mechanism is what our system aims to fix.

The very nature of this problem seems a very good candidate for the use of IoT technology in order to automate the device inspection process. Using sensors to monitor physical aspects of the neutralization system, we can, not only, extract key information about the device, but also notice the owner of the device when the next maintenance will be needed.

1.2 Objectives

The project aims to leverage **Internet of Things (IoT)** [5] technology to automate the monitoring and maintenance of the filtration system.

By integrating a microcontroller unit (MCU)[6] with sensors, the system will gather data on key filter variables and transmit it to a central server.

Establish a robust central server as the backbone of the system: One of the key objectives of our project is to develop one robust central server that serves as the backbone of the entire system. This central server will play a crucial role in storing, processing, and managing the all collected data from all the filtration systems. It will have multiple functionalities:

- **Data storage and analysis:** The central server will include data storage mechanisms to efficiently store and manage the collected data. It will also process and analyse the data, extracting meaningful insights and patterns;
- **Broker functionality:** The central server will incorporate a broker component that acts as an intermediary between the server and the microcontroller unit (MCU). This broker facilitates seamless communication and data exchange between the MCU and the server, ensuring reliable and efficient transmission of information;
- **Web API for system integration:** To enable seamless integration with other systems and applications, the central server will expose a *Web API*[7]. This API will allow different systems to consume the collected data and access specific functionalities provided by the server. It will provide a standardized interface for easy and secure interaction with the system.

/Enable remote monitoring and data visualization: The development of a user-friendly web application serves as a crucial objective in this project. It will allow users to remotely access and visualize the collected data from the filtration system. This interface will provide easy interaction and comprehensive insights into the system's performance, enabling efficient decision-making and proactive maintenance.

To ensure the project's value, competitiveness, and appeal, several requirements need to be fulfilled. Here are the guidelines we have agreed upon to steer the project in the right direction:

1. Cost-effective integration of the microcontroller unit into the existing product.
2. Prolonged battery life for minimal maintenance requirements.

3. Comprehensive documentation for user-friendly operation and understanding.
4. Clean and maintainable code for easy unit testing.
5. pH data reading from the filter and sharing it with a server.
6. Notification to the filter owner or manager if any value exceeds a set threshold, i.e. proactive maintenance.
7. Intuitive and user-friendly website design for easy navigation and access to information.

1.3 Document Organization

This section provides an overview of the organization of the chapters of this report, outlining the main topics and sections covered.

The document organization ensures a logical flow of information and facilitates understanding of the project's scope, objectives, and findings.

In the **first chapter**, it is provided information about the problem we are trying to solve and how our project solution is an interesting way to approach it, we also state the overall project objectives and the requirements we proposed to follow.

The **second chapter** provides important contextual information. It lays the groundwork for comprehending the project's strategies and technologies used, as well the decisions taken to choose each element of the system.

The purpose of the **third chapter** is to provide an in-depth explanation of the system architecture, dividing the problem in different modules and make a clear explanation of how each part of the problem was approached and the challenges we encountered during the implementation process.

The **conclusion** chapter outlines the most important aspects of solution and as well as the most difficult changes we had to face during the duration of the project.

2

Background

In this chapter will be provided the necessary information for the understanding of the core elements and technologies of our proposed architecture, as well as the analysis of the different alternatives for each component selected. Additionally, the chapter will offer general insights into IoT (Internet of Things) technology, providing a broader understanding of its principles and applications.

2.1 IoT Systems

The Internet of Things (IoT) refers to the vast network of interconnected devices, objects, and systems that are capable of collecting, exchanging, and sharing data over the Internet [5]. Regarding the physical components, a crucial aspect of IoT systems is the deployment of sensors, as they play a vital role in gathering data from the surrounding environment. The collected data is transmitted to a central system or cloud platform for processing, storage, and analysis. That system is responsible to centralize data management, process incoming sensor data, and facilitate communication and coordination between connected devices. With this architecture we can create a system that enable seamless communication between physical devices, collect and analyse data from those devices, and utilize the insights gained to improve efficiency, automation, and decision-making in various domains such as healthcare, transportation, manufacturing among others.

2.1.1 Sensors

To achieve our solution, we needed to implement several sensors. The **pH measurement** circuit contains a pH probe that oscillates between positive and negative voltages, representing pH values. This sensor plays a crucial role in our system as it is responsible for analysing the pH of the liquid [8]. It serves as a tool for evaluating the productivity of the neutralization mechanism, providing valuable insights into its performance.

The **ambient temperature and humidity measurements** were carried out by sensor DHT11 [9]. This sensor is also factory calibrated, making it easy to interface with other microcontrollers. Moreover, it is capable of measuring temperatures

ranging from 0 °C to 50 °C and humidity from 20% to 90%, providing an accuracy of ± 1 °C and $\pm 1\%$. The values of the temperature and humidity are outputted as serial data by the sensor's 8-bit microcontroller[10]. This sensor offers valuable insights into the operational environment of the mechanism. Its advantage lies in its ability to gather extensive data regarding the working environment, which can be utilized by manufacturers to enhance the product further.

The DHT11 Sensor is factory calibrated and outputs serial data, making it highly easy to set up. The connection diagram for this sensor is shown below.

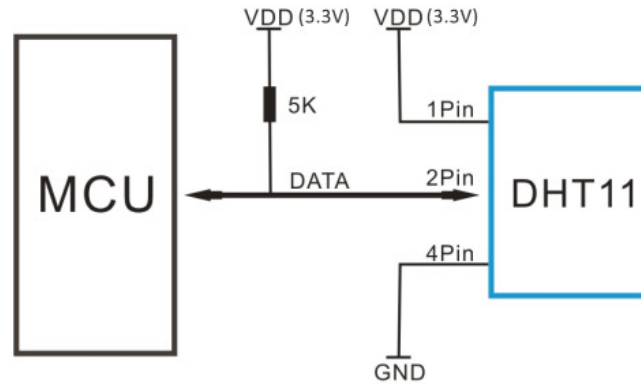


Figure 2.1: DHT11 Sensor Circuit

As shown in the Figure 2.1, the data pin is connected to an I/O pin of the MCU and a 5K Ohm pull-up resistor is used. This data pin outputs the value of both temperature and humidity as serial data.

The output given out by the data pin will be in the order of 8bit humidity integer data + 8bit the Humidity decimal data + 8 bit temperature integer data + 8bit fractional temperature data + 8 bit parity bit. To request the *DHT11* module to send data, the I/O pin has to be momentarily made low and then held high, as shown in the Figure 2.2.

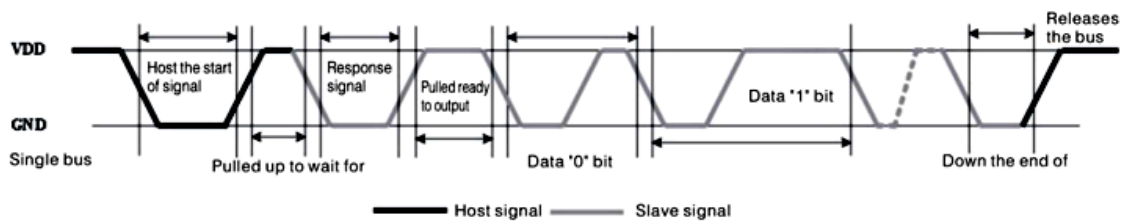


Figure 2.2: DHT11 Data Timing Diagram

To detect **water leaks**, in our scenario, small water droplets near the equipment's outlet, we chose a simple water detector sensor that would be attached to the valve of the mechanism, it has a 3-pin module that outputs an analogue signal that indicates the approximate depth of water submersion. Overall, the more water it gets, the more conductivity and the voltage increases. When the sensor is dry, it outputs zero voltage. The primary consideration was the sensor's wide coverage and easy placement near the mechanism's exit, enabling effective detection of water leaks.

2.1.2 Embedded Systems

Embedded systems are compact and energy-efficient computers that are integrated into various mechanical or electrical systems. These systems are characterized by their low cost and power consumption. Typically, an embedded system consists of a processor, power supply, memory, and communication ports. These communication ports facilitate the exchange of data between the processor and peripheral devices[11].

MCU (microcontroller unit)

The primary requirement for selecting the MCU was its compatibility with integrating a Wi-Fi module or having built-in Wi-Fi capabilities, as the device needed to transmit data over the internet.

We chose to utilize Wi-Fi instead of Bluetooth as our wireless communication protocol due to several advantages. Firstly, Wi-Fi offers a greater range than Bluetooth, ensuring reliable communication between the pH filter units and the access point even when they are positioned at considerable distances from each other. This is particularly advantageous if the pH filter units are spread across a large area or if there are obstacles that could potentially disrupt signal transmission. Additionally, Wi-Fi networks are more prevalent and established in various environments, including homes, offices, and public spaces. This widespread availability makes Wi-Fi a suitable choice for our project.

One requirement of this project was the overall low cost of the system, since we are imitating an industrial project as closely as possible. The battery consumption of the selected microcontroller was another requirement, as the device's ability to successfully automate the water inspection process relied on its ability to operate for an extended period of time. During this phase, we encountered a constraint related to the time-sensitive nature of the project, which limited our choice of microcontrollers (MCUs) available in the market since they could not be obtained within the required timeframe.

The main options we took into account were part of the *espressif* family: the *ESP32 - S3*, *ESP32 - S2*, *ESP8266EX*, *ESP32 - C3* microcontrollers. However, we also considered MCUs from other manufacturers, such as the *MSP430FR413x* from *Texas Instruments*. The following tables reflect the main criteria under analysis:

During the evaluation process, we primarily considered three key factors for selecting the board. First, the presence of a built-in Wi-Fi module was crucial as the board needed to connect to Wi-Fi networks. This feature ensured seamless wireless connectivity for our project.

Secondly, we paid close attention to the number of *General Purpose Input/Output (GPIO) pins* available on the board. This factor was vital for the future expandability of the project, allowing us to connect a sufficient number of sensors and peripherals as needed.

Lastly, our main emphasis was on the power consumption during the *deep sleep* mode, which serves as the primary operational state of the board for extended

Microcontroller Comparison				
Model	ESP32 - S3	ESP32 - S2	ESP8266EX	ESP32 - C3
Active Mode (mA)	91 - 340	68 - 310	56 - 170	95-240
Modem Sleep (mA)	13.2 - 107.9	10.5 - 32.0	15	18 -28
Light Sleep (uA)	240	750	900	130
Deep Sleep (uA)	7-8	20 -190	20	5
Power Off (uA)	1	1	0.5	1
Wifi	yes	yes	yes	yes
Bluetooth	yes	yes	yes	yes
GPIO	34	43	17	22
CPU	dual-core 160MHz to 240 MHz	single-core 240 MHz	single-core processor that runs at 80/160 MHz	single-core 160 MHz
Price (€)	14.1	7.52	-----	8.46

Table 2.1: Comparison Between different MCUs from espressif family, prices consulted from Mouser Online Store at 10/03/2023, URL: <https://pt.mouser.com/>

Microcontroller Comparison	
Model	MSP430FR413x
Active Mode (mA)	126 uA / MHz
LPM0	158 - 427 uA (20 uA / MHz)
LPM3	1.25 - 1.99 uA, 25°C (1.2uA)
LPM4	0.57 - 0.75 uA, 25°C (0.6uA without SVS)
LPM3.5	0.70 - 1.25 uA, 25°C (0.77uA with RTC only)
LPM4.5	0.013 - 0.375 uA, 25°C
SHUTDOWN	13 nA
Wifi	-----
Bluetooth	-----
GPIO	14.1
Price (€)	17.49

Table 2.2: MSP430FR413x analysis, price consulted from Mouser Online Store at 10/03/2023, URL: <https://pt.mouser.com/>

durations[12]. This mode was prioritized due to the device’s minimal data transmission activities. A comprehensive explanation of how we effectively utilized this mode will be provided in the architecture chapter of this report3. Additionally, we considered the power consumption during active mode, which would occur when the board collected data from the sensors and transmitted it.

By carefully considering these factors, we were able to make an informed decision regarding the selection of the board for our project.

The **ESP32 - S3** is a high-end MCU that offers an impressive array of features, including a built-in Wi-Fi module, low power consumption, and a powerful processor [13]. However, it comes with a relatively high price tag. While the ESP32 - S3 provides

a wide range of possibilities for our project, one of the requirements is to maintain awareness of the final product's cost-effectiveness. Considering the overall economy of the project, we opted against choosing the ESP32 - S3 due to its higher cost.

The ESP32-S2 is a versatile microcontroller unit (MCU) that offers a wide range of capabilities suitable for various projects. Despite having a slightly less powerful processor compared to some other options like the ESP32 - S3, the ESP32-S2 provides sufficient processing power for the intended tasks. One notable feature of the ESP32-S2 is its built-in Wi-Fi module. The MCU offers the advantage of low power consumption, making it an excellent choice for energy-efficient applications. Furthermore, the ESP32-S2 offers a cost-effective solution without compromising essential features. Overall, the ESP32-S2 is a reliable and capable MCU that strikes a balance between performance, power efficiency, and affordability [14].

The **ESP8266EX** board stands out as a choice for our project due to its exceptionally low power consumption, despite having a slightly less powerful processor compared to the other options[15]. Its built-in Wi-Fi capability is also a valuable feature. One drawback of the ESP8266EX is its low number of GPIO (General Purpose Input/Output) pins comparing to the other options. This limitation could potentially restrict the number of sensors that can be added to the project, thereby limiting its expandability in the future. Although we were unable to select this MCU for our project at this stage due to stock shortage, it is worth mentioning as a notable competitor that we considered.

The **ESP32 - C3** is a noteworthy contender for our project. It has a combination of features, including low power consumption, a built-in Wi-Fi module, and similar characteristics to the ESP32-S2 [16]. Although it runs on a slightly less powerful CPU, it delivers excellent deep sleep consumption. However, it is important to note that the ESP32-C3 is slightly more expensive compared to the ESP32-S2. Despite this cost difference, its overall performance and feature set make it a strong candidate for consideration in our project.

The **MSP430FR413x** microcontrollers has superior power consumption, making it an energy-efficient choice. However, it lacks a built-in Wi-Fi module, which presents a challenge [17]. To incorporate Wi-Fi functionality, an additional purchase and installation would be required, adding complexity and potentially increasing costs. Considering our project's time constraints and the readily available access to the ESP32-S2, it was important to choose a solution that offered convenience and efficiency.

Considering the project's time constraints and the availability of the ESP32-S2, along with the cost and additional requirements associated with the MSP430FR413x line-up, we have opted to prioritize convenience. As a result, we have chosen the ESP32-S2 as the preferred platform for our project.

In conclusion, the ESP32-S2 emerged as the most suitable and practical option for our project, ensuring smooth implementation without compromising essential features. Figure 2.3 shows the version of the chosen microcontroller.

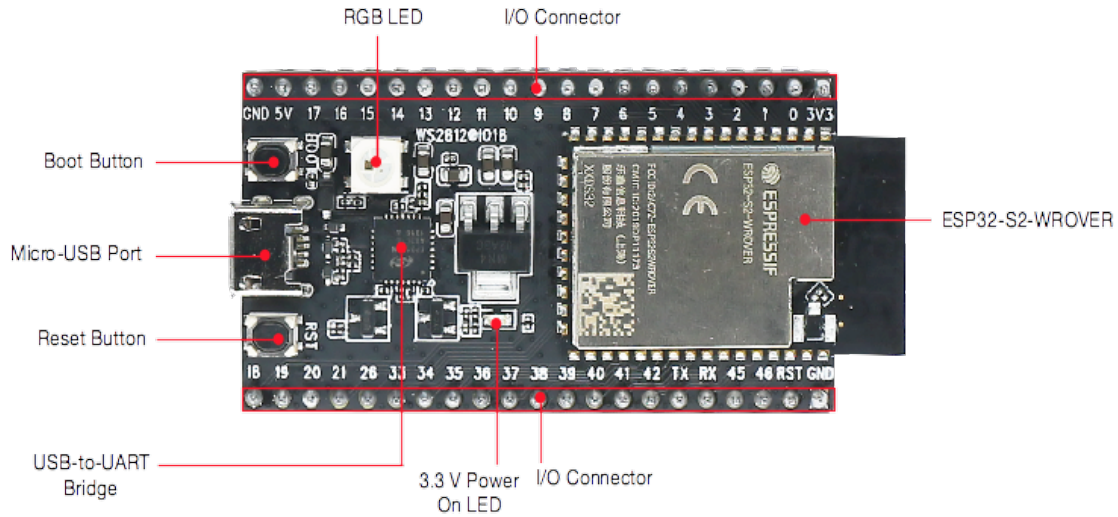


Figure 2.3: Espressif ESP32-S2-SAOLA-1M (WROOM)

Framework

We made the decision to utilize the **ESP-IDF** (Espressif IoT Development Framework) framework, the development framework for Espressif SoCs supported on Windows, Linux and macOS, instead of the more commonly used and less versatile Arduino and PlatformIO frameworks for programming the MCU.

This choice was motivated by the project's requirement for an economic power device, as highlighted in the section 1.2. The ESP-IDF framework provides us with direct access to the underlying hardware and peripherals, giving us more control and flexibility over the microcontroller. This allows us to optimize performance in a variety of tasks.

Another advantage is the documentation freely available. On the other hand, many tasks, for instance connecting to the Wi-Fi or interacting with the sensors, becomes a more complex task due to the non-existent abstraction that other frameworks provide.

2.1.3 Backend

In an IoT system, the backend serves as the central component responsible for managing and processing the data collected from connected devices. The main functions of the backend in an IoT system include:

Technology Stack

In the backend development of our IoT project, we adopted a technology stack consisting of the Spring framework and the Kotlin programming language. In the backend development of our IoT project, we adopted a robust technology stack consisting of the Spring framework and the Kotlin programming language.

The **Spring framework** served as the foundation of our backend implementation. We leveraged its comprehensive features, MVC architecture, transaction man-

agement, and seamless integration with other components[18].

The use of this framework provided us with a scalable and flexible infrastructure for building our backend services.

Kotlin, a modern and expressive programming language, was chosen as the primary language for our backend development [19]. This decision was based on a thorough evaluation of its features and benefits. Kotlin's concise syntax enabled us to write code that is more succinct and easier to comprehend, and, simultaneously, not as verbose as Java. The built-in null safety feature significantly reduced the occurrence of null pointer exceptions, a common source of bugs in other languages. Additionally, the rich standard library provided a wide range of utility functions and classes, enabling us to implement complex functionalities more efficiently.

Due to our prior exposure to and experience with both of these technologies throughout our course, we have developed a strong familiarity with them. This familiarity significantly influenced our decision to incorporate Spring and Kotlin into our project's tech stack.

Database

To store the data, we decided to separate the data in two sections: The **non-time-dependent data**, which encompasses the portion of data within a database that remains constant and undergoes minimal or no changes over time or during system operation, such as information about the users, error logs and associated devices. The main considerations in selecting a technology for storing this type of data were scalability and support for intermediate querying capabilities. The open-source relational database *PostgreSQL* emerged as an ideal choice that fulfilled these criteria. Furthermore, our prior exposure to with this technology during our course expedited the development process of this module, enhancing overall time efficiency.

The **sensor data**, which is derived from the continuous collection of data by the device's sensors, such as pH values, is characterized by its time-stamped nature and represents a continuous stream of values or events. With this design in mind, we chose to utilize a *time series* database to store all the data related to the sensor records.

Time series databases are specifically designed to store and retrieve data records associated with timestamps (time series data), due to the uniformity of time series data, specialized compression algorithms can provide improvements over regular compression algorithms designed to work on less uniform data. They offer faster querying and can handle large volumes of data[20][21][22]. However, it's important to note that these databases may not perform as effectively in domains that do not primarily work with time series data.

For our project, we made the decision to utilize the *InfluxDB* time series database, a widely recognized and widely used time series database solution [23]. *InfluxDB* offers a range of advantages that we previously discussed, and it seamlessly integrates with Kotlin, the server-side language we have adopted.

2.1.4 Communication protocols

After careful consideration, we opted for the **MQTT** protocol to allow the communication between the device and the server. MQTT is extensively utilized in the IoT industry due to its lightweight and efficient nature. This efficiency enables devices to conserve energy more effectively in comparison to other protocols such as HTTP[24][25]. MQTT achieves this by employing a smaller packet size and lower overhead than its counterparts. To allow the communication between the web app client, we used the HTTP protocol. Our application provides an API (Application Programming Interface) that allow other applications or services to interact with them. APIs are typically based on HTTP and use various HTTP methods (such as GET, POST, PUT, DELETE) to perform operations like retrieving data, creating resources, updating resources, or deleting resources. API clients make HTTP requests to these endpoints to interact with the web application and exchange data.

Broker

One of the fundamental components of the solution is the presence of a broker. On an IoT system, data is consistently transmitted from the devices to the back-end server. Therefore, it is crucial to select a communication protocol that is both lightweight and offers a dependable communication channel.

The role of the broker in the MQTT protocol is to serve as an intermediary, facilitating the transmission of data from publishing clients to subscribing clients. In the specific context of our project, the publishing clients are represented by various MCUs (Microcontroller Units), while the subscribing client is the central server tasked with analysing the data.

To implement the broker, we have decided to utilize the free version of the HiveMQ CE. This selection aligns with our requirements, particularly our focus on message security.

The *HiveMQ* broker supports TLS/SSL encryption, ensuring secure communication between *HiveMQ* and MQTT clients (both publishers and subscribers). Additionally, it provides robust support for handling substantial amounts of data and is compatible with Kotlin. While the free version of *HiveMQ* does have limitations, such as a maximum allowance of 100 devices, it is deemed sufficient for the scope of our project.

MCU Configuration

To facilitate interaction between the MCU and the equipment owner, we decided to adopt a solution provided by the manufacturer of the ESP32-S2. The ESP Touch protocol is a wireless communication protocol developed by *Espressif Systems* specifically for their *ESP8266* and *ESP32* series of Wi-Fi enabled microcontrollers. The free ESP Touch mobile app uses this protocol, to connect to the ESP device, and thus, transmitting, not only the Wi-Fi configuration, but also a string of bytes of custom data. With the objective of transmitting the data, the protocol utilizes a smart config-

uration mechanism where a user's smartphone or device with the ESP-Touch app establishes a Wi-Fi connection with the ESP32-S2 module. the packets are broadcasted by the smartphone, allowing the ESP32-S2 to receive and decode them.

Due to our selection of the ESP32-S2 microcontroller, we have decided to incorporate this protocol into our system for MCU configuration purposes.

2.1.5 Frontend

The Frontend refers to the client-side part of a software application that users interact with directly. It is responsible for presenting the user interface (UI) and handling user interactions, such as inputting data, submitting forms, and receiving and displaying information. In our project, we have developed a user-friendly frontend interface that offers users visualizations of the collected sensor data. This interface allows users to interact with the data and gain meaningful insights in an aesthetically pleasing manner.

To develop the frontend, we had to select a web app client framework. The web app client, which comprises the code and resources delivered to the user's browser and executed there, is responsible for creating the user interface and facilitating user interactions. It serves as the practical realization of the frontend using web technologies. To select the appropriate technology, we evaluated several options and conducted thorough analysis.

Given our previous experience with **React**, a widely adopted framework for building front-end applications, we made the deliberate choice to utilize it in our project. The majority of the website's styling was accomplished using pure CSS.

To streamline our development process, we integrated **Webpack**, a free and open-source module bundler for JavaScript. We also took advantage of TypeScript, which provides benefits such as type checking, enhanced code editor support, and improved code documentation. Webpack was configured to employ a TypeScript loader responsible for transpiling TypeScript files into JavaScript, thereby ensuring compatibility with various browsers[26].

By adopting these technologies and tools, we successfully constructed a visually appealing and interactive front-end interface while leveraging the advantages of React, React Bootstrap, Webpack, and TypeScript.

2.2 Existing solutions

CropX specializes in delivering IoT-based soil monitoring systems that incorporate pH sensors. Their extensive experience in the market has enabled them to provide real-time data on soil conditions, empowering farmers to make informed decisions regarding irrigation and nutrient management. The fundamental principles of this industrial solution align closely with our own project, as both aim to equip users with valuable insights derived from data, thereby facilitating better decision-making processes. They also provide an intuitive UI (User Interface) for the user to analyse the data[27].

In 2022, a project was developed by Bernardo Marques, under the guidance of Professor Rui Duarte. This project is called **SmartDoorLock-IoT**. It consists of a system composed by a microcontroller, in particular an ESP32-S2, that hosts a socket based server and an Android application that interacts with the MCU/Server to lock and unlock a door. Behind the curtains, the system uses security features, like encryption through symmetric and asymmetric keys, Hash, MAC, databases and Wi-Fi communication. Using an IoT device, the system is able to take advantage of features only possible with microcontrollers, like the ultra low power mode, available in the ESP32-S2, which is very important in systems of this kind. Overall, this project/system takes advantage of the IoT technology to implement a secure IoT system capable of simplifying the life of a human being.

2.3 Summary

This chapter provides the background information in order to gain a comprehensive understanding of the core elements and technologies employed in the proposed architecture. It discusses the various options considered, and the decisions made when selecting specific technologies that align with the project requirements. Furthermore, the chapter offers explanations on the purpose of each chosen technology and highlights its relevance to the overall project objectives.

3

Proposed IoT System Architecture

In this chapter, the proposed system architecture will be presented, starting with the system high level view. This means defining each module and how they all interact with each other to compose the final solution. Then, for each component presented, the details of its implementation will be addressed. This way, the reader can, not only, understand the proposed system architecture without relying on the implementation details, but also deep dive into each component specific implementation.

3.1 System Overview

It is possible to divide the system into the following sections:

1. Factory: Collects the data from the neutralization filter;
2. Backend: Receives and processes the data received from the MCU;
3. Web Application: Allows the user to analyse relevant filter-collected data, and manage devices.

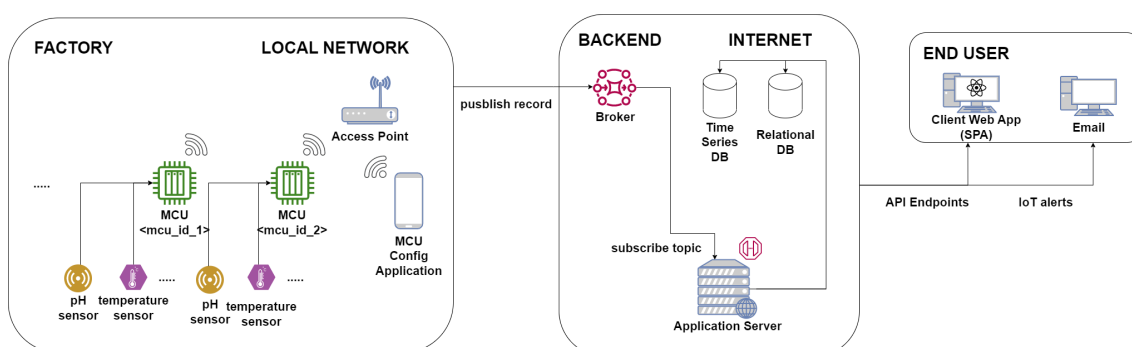


Figure 3.1: Main system architecture

As shown in Figure 3.1, data flows from left to right. All begins in the MCU collecting data - *sensor records* - and then, sending the sensor data to the *Backend Broker* module. The Backend is composed by the Broker, which is the middleman between the *IoT Node* and the application server. The latter receives the sensor data from the

broker to process it, accordingly. Eventually, the application server might send an alert email to the manager of the neutralization mechanism if the data indicates an abnormal condition. The server also provides an API that allows other applications, such as the web application described in this report, to interact with the system. This API enables functionalities like querying available sensor data, creating users, and adding devices.

3.2 IoT Node

Each client utilizing the system outlined in this report will have a replicated component consisting of four subcomponents, as presented in the figure 3.2.

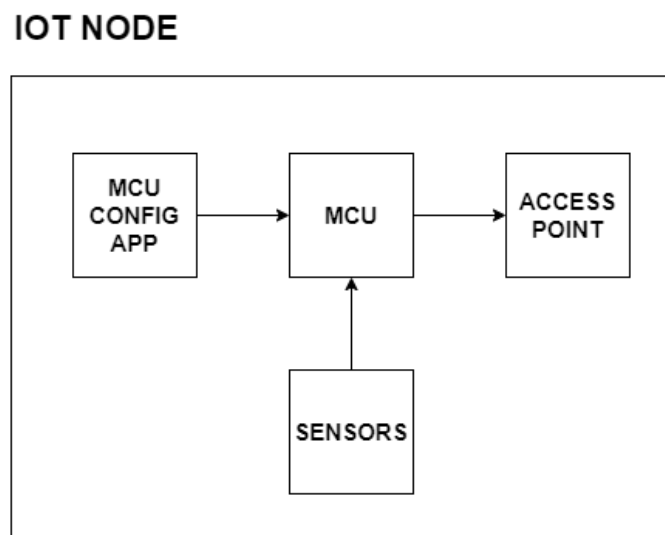


Figure 3.2: IoT Node Module Architecture

The subcomponents have the following functions:

1. Sensors: collect neutralization filter data;
2. Microcontroller (MCU): coordinate when to collect and send the data to the Backend;
3. Access Point (AP) to enable the MCU to connect to the internet;
4. Android app to configure the MCU.

With the objective of setting up the MCU for the first time, the owner of the mechanism that is associated with the MCU interacts with an android app provided by **Expressif** called **Esp Touch** to configure the MCU. This process consists in setting up a device identifier provided by the web application developed by us, in which the user must create an account beforehand, and setting the Wi-Fi network credentials. The Esp Touch application will be responsible for transmitting this data to the MCU. After, the MCU will instruct the sensors to collect data, and send it to the backend, through the access point.

3.2.1 MCU

In the majority of cases, the MCU operates in a low-power mode known as "deep-sleep." In this mode, the CPUs, most of the RAM, and digital peripherals that are clocked from the APB.CLK (Advanced Peripheral Bus Clock) are powered off [12]. This low-power mode is maintained throughout most of the day, as continuous readings are not required due to the nature of the neutralization mechanism. Changes in conditions occur over longer periods of time. Given the significance of power consumption, every second that the MCU remains active has a notable impact.

The MCU is susceptible to hard resets, which can occur when it is in a sleep state. During a hard reset, the MCU's unique identifier and Wi-Fi credentials are erased, leading to a loss of configuration. In such cases, or when the device has not been initially configured, it will remain idle, waiting for proper configuration to be provided.

Upon waking up, the device performs an analysis to determine the cause of its wake-up. If the device has been powered on, it proceeds to calculate the calibration times for each sensor. These calibration times denote the duration that each sensor needs to wait after being turned on before initiating the measurement process. A detailed explanation of this process is provided later in the section 3.2.4.

Typically, the ESP will re-start execution after aborting from an exception that was not handled. When this happens, the device reports the error cause to the backend, and goes into deep-sleep mode again. It is important to collect all the error logs since they provide relevant information about the state of the equipment. After an error occurs in the system, the error information will be sent to the backend to be stored and later analysed.

The standard programmed behaviour of the MCU is divided into two main cycles:

- The MCU will use all sensors to make readings, including the pH sensor, the DHT11 and the water sensor. First, all sensors are used to collect readings and only then, the Wi-Fi is powered on. This cycle is the one that consumes the most energy, since the Wi-Fi will be used to send the readings to the broker. In case there is an issue and the readings are not sent over the network, they will be stored.
- In this cycle, the MCU is responsible for monitoring the neutralization system for water leakage. It accomplishes this by using a water sensor to detect the presence of water and subsequently sending an alarm to the backend if necessary. In the event that water is detected, the MCU utilizes Wi-Fi to transmit an alert to the broker. It also takes the opportunity to check if any sensor readings from the previous cycle were not successfully sent due to any issues. If there are pending readings, they are transmitted during this cycle. However, if no water is detected and there are no pending readings, the MCU will disregard the readings. This cycle typically consumes less energy compared to the previous cycle, since Wi-Fi is not actively used when no water leakage is detected. Additionally, the calibration time for the water sensor is relatively shorter than that of

the other sensors. As a result, this cycle can be repeated more frequently without significantly impacting energy consumption, thereby enabling better control over potential water leakage events.

It is worth mentioning that in situations where Wi-Fi is unavailable, the sensor readings will be stored in the device to be transmitted during the next internet connection. If the device wakes up and still does not have access to Wi-Fi, the latest reading will be stored while the previous reading will be discarded. This approach was chosen to optimize battery usage by sending smaller payloads. Since the primary objective of the system is to actively alert users, it is deemed that the most recent reading holds the utmost importance, rendering previous readings obsolete.

3.2.2 Sensor Modules

The primary objective of this project is to develop a system that automates the inspection process for potential issues in a neutralization mechanism, but is also interesting to develop a system that is capable of providing us with meaningful insights, regarding the utilization of the filter. Taking that into consideration we opted for the following sensors:

1. Initial pH sensor to obtain the solution pH in the beginning of the filter system (before being neutralized)
2. Final pH sensor to obtain the solution pH in the end of the filter system (after being neutralized);
3. pH sensor to obtain the pH of the solution present in the filter system (before being neutralized);
4. Ambient temperature sensor;
5. Air humidity sensor;
6. Water flow sensor to estimate the amount of water that flows in the system;
7. Water detection sensor to alert if water has exited the system through a place it is not supposed to.

The purpose behind the installation of two pH sensors, one at the filter's entrance and another at the exit, is to gather data on the product's efficiency. By comparing the pH levels, the manufacturers of the equipment can estimate the effectiveness of the system.

During the project's planning phase, we identified a potential constraint regarding the availability of physical sensors. This raised concerns about our ability to conduct comprehensive system testing. To overcome this challenge, we devised a solution by implementing mock implementations for the "Sensor Reader" modules. These mock implementations were designed to simulate the actual sensor readings, enabling us to evaluate the system's functionality in the absence of physical sensors.

Consequently, for each sensor reader component, two implementations were developed: a real implementation that interacts directly with the physical sensor, and a mock implementation that simulates the sensor procedure. This approach provided flexibility and independence from the availability of the physical sensors,

allowing us to continue testing and developing various components of the system regardless of the sensor's presence.

3.2.3 Sensor readings - Parallelized vs Serialized

Choosing the approach for reading from the sensors is a crucial consideration. One option is to start all the sensors simultaneously and parallelize the reading process, while the other option is to read each sensor individually and serialize the reading operations.

If we opted for parallelized reading, the main advantage is that the microcontroller would read from all the sensors at the same time. This means that the total active time of the microcontroller would be determined by the sensor that takes the longest to complete its operations. Consequently, the microcontroller can enter sleep mode faster, conserving more power.

However, it is important to consider the trade-offs of this approach. Parallelized reading may introduce challenges such as unstable or unreliable operation of the sensor, since all the sensors are activated at the same time, they may collectively draw a higher current, potentially exceeding the power supply capacity or putting unnecessary strain on the power source. It would also require more complex synchronization mechanisms.

On the other hand, serializing the reading operations by individually starting and reading each sensor can help mitigate these challenges. It simplifies the hardware design and reduces the likelihood of interference or conflicts between the sensors. Although this approach may require more time for data acquisition compared to parallelized reading, it offers greater control over power consumption, since it allows for selective activation of individual sensors and can ensure more reliable and accurate measurements from each sensor due to the lack of conflicts between the sensors. After carefully evaluating these advantages and disadvantages, we opted for serializing the readings of the sensors.

3.2.4 Sensor Calibration

An important aspect we need to consider is the sensor calibration, which involves determining the required waiting time for a sensor after it is turned on before it can start collecting accurate data from the environment. This waiting time is commonly referred to as the stabilization time. During this period, the sensor adjusts and stabilizes its internal components to ensure reliable and consistent measurements.

To address the variability of sensor stabilization time in different environments, we developed a method to calculate the required waiting time for each sensor independently, regardless of its specific location.

The approach we employed to determine sensor stabilization is based on the concept of standard deviation. Standard deviation measures the average deviation of individual values from the mean. A higher standard deviation indicates that the values are more dispersed and further from the mean, indicating that the sensor is not yet stabilized. Conversely, a lower standard deviation suggests that the values

are closely clustered around the mean, indicating that the sensor readings are approaching stability.

To establish our strategy, our initial goal was to determine the standard deviation of the sensor readings when it reaches a stable state. This required conducting experimental analysis of the sensor's behaviour. We powered on the sensor and recorded the corresponding ADC values it produced every second. By closely observing the readings, we identified the point at which the sensor reached stability, characterized by low variability in its outputs, indicated by a low standard deviation.

To precisely determine when the sensor reaches a stabilized state based on the experimental data, we employed an algorithm known as the sliding window. This algorithm involves creating a virtual window of a fixed size, denoted as 'x', which slides through the array of ADC values obtained from the sensor. At each iteration, the algorithm calculates the standard deviation of the values within the window.

By applying the sliding window algorithm, we can track the changes in the standard deviation as the window moves through the ADC value array. Initially, when the sensor is not yet stabilized, the standard deviation will be relatively high due to the variability in the readings. However, as the window progresses and includes more stable readings, the standard deviation will gradually decrease. In Figure 3.3 we can see a visual representation of oh this algorithm works.

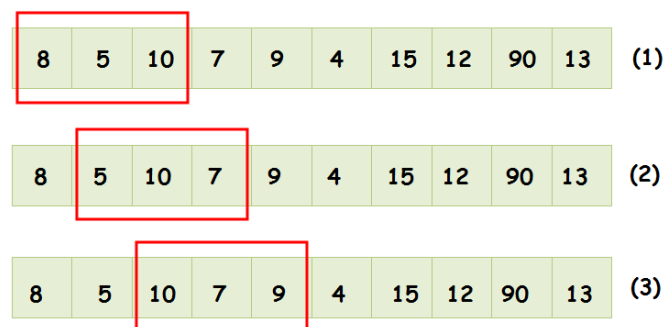


Figure 3.3: Sliding window algorithm

Once we observe that the standard deviation consistently remains low, we can infer that this particular standard deviation corresponds to a stabilized sensor. However, a crucial question remains: what should be the size of the data window that we analyse?

The choice of the window size in the sliding window algorithm is a crucial factor in accurately assessing the stabilization of the sensor. If the window size is too small, it may not capture sufficient data points to provide a representative estimate of variability. Consequently, the calculated standard deviation within such a small interval might be misleading, potentially indicating stabilization at an early stage. However, this would not be a reliable indication, since the limited data within the small window may not be truly representative of the overall behaviour of the sensor.

On the other hand, if a larger window size is used, the calculation of the standard deviation may take longer due to the increased number of data points being considered. It becomes more challenging to achieve a low variability in a larger ar-

ray of data, prolonging the time required for the sensor to exhibit stable readings. This implies that a longer waiting time would be necessary for the sensor to reach a desirable level of variability within the analysed data window.

Therefore, finding an optimal window size is crucial to strike a balance between capturing sufficient data for reliable analysis and minimizing the waiting time for stabilization.

Taking as example the pH sensor:

Through empirical observation under experimental conditions, we determined that the pH sensor exhibits a stable state when the standard deviation of its ADC values approaches a value of approximately 2.

To determine the appropriate window size for calculating the standard deviation and assessing sensor stabilization, we conducted experiments using various window sizes. By analysing the standard deviations of the ADC values received by the MCU, we aimed to identify the point at which the values reached a standard deviation of 2, indicating sensor stabilization. Through this iterative process, we evaluated the effects of different window sizes on the accuracy and timeliness of detecting sensor stabilization.

The following graphs depict the standard deviation over time obtained using two different window sizes, 3 and 30. In each graph, the standard deviation values are plotted, and the line turns red when the sensor is considered stabilized, meaning that the standard deviation is equal to or less than 2. These graphs provide a visual representation of how the sensor's stability evolves over time for different window sizes, allowing us to observe the importance of choosing an indicate size for the window.

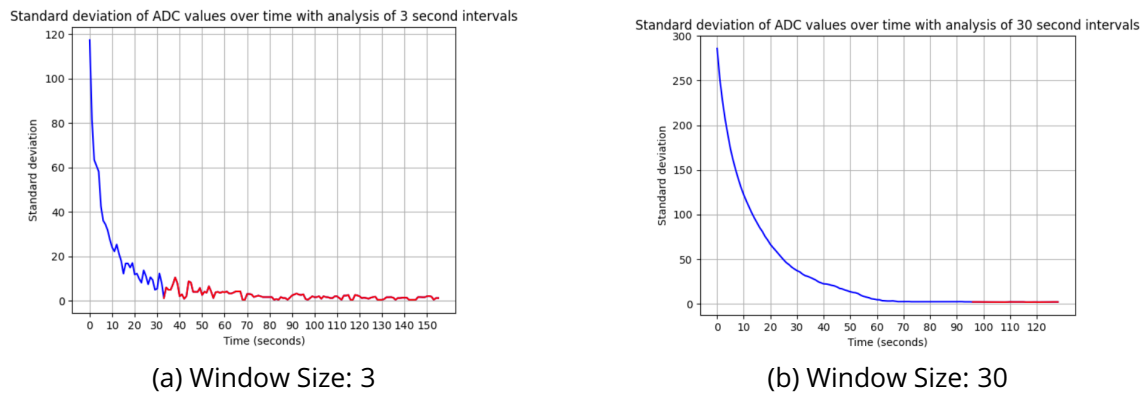


Figure 3.4: Standard Deviation over Time

By analysing these graphics, we can visually observe the trend that larger window size (30) require a longer time than is necessary to time for stabilization, while smaller window size (3) result in faster stabilization. However, it's important to note that with smaller intervals, the sensor may not be fully stabilized at that specific time, as we can observe in Figure 3.4a.

Knowing this propriety, we tested different window sizes and registered the respective time for stabilization, comparing to the window size, as shown in the Table 3.1.

After conducting the analysis, we have chosen a window size of 5. This decision

Table 3.1: Window Size vs. Time to Stabilize

Window Size	Time to Stabilize (seconds)
3	33
4	42
5	74
10	75
15	73
20	71
25	99
30	96
35	96
40	93
45	93
50	95

Standard deviation of ADC values over time with analysis of 5 second intervals

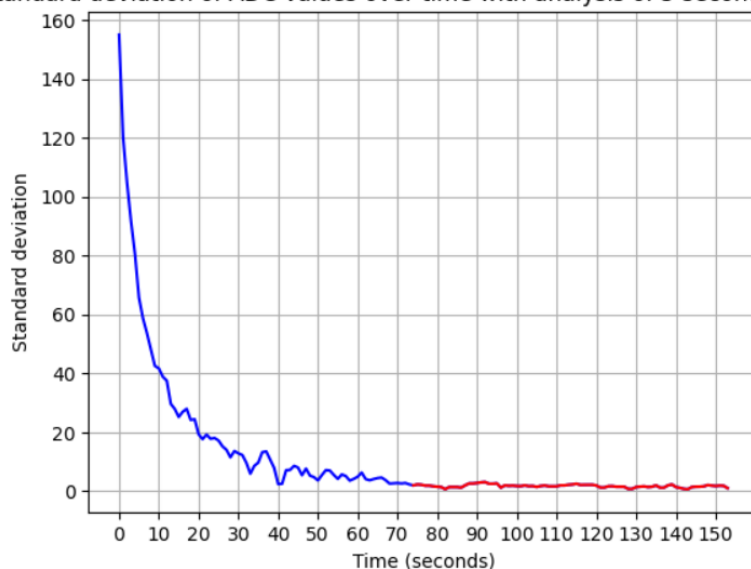


Figure 3.5: Standard Deviation over Time

was made based on the observation that window sizes 10, 15, and 20 exhibited similar values for the time to stabilize. By visually inspecting the graph, we can confirm that the values tend to stabilize at around 74 seconds. In our experiments, we found that a window size of 5 strikes a balance between not causing excessive delay in the stabilization process and still being representative of the system's variability. We can visually observe the stabilization of the sensor in Figure 3.4a, where the red line indicates the point at which the standard deviation of a five-element interval reaches or falls below 2. This signifies that the sensor has achieved stability.

With a fixed window size of 5, the calibration process of the pH sensor commences is the following: When the MCU is powered on, it initiates the calibration process by sequentially activating each sensor. During the calibration of the pH sensor, it collects readings from the environment at a frequency of one reading per second. To calculate the standard deviation, a sliding window algorithm is employed. Once the standard deviation of a 5-element interval reaches a value that is equal to

or lower than the reference value, it indicates that the sensor has stabilized. Therefore, the time at which this stabilization occurs can be considered as the required time for the sensor to stabilize. This time will be stored persistently in the device for future readings.

To address the variation in environmental conditions and account for different scenarios, we adopted a more conservative approach in determining the stabilization of the sensor. We increased the reference value of the standard deviation by fifty percent. Although this value is significantly higher than the standard deviation in stabilized conditions, it provides some tolerance to accommodate the variability of different environments. In a real-world application scenario, we consider a standard deviation of 3 as an indicator of sensor stabilization. However, it's important to note that this value may not be entirely accurate if the environmental differences in the filter mechanism significantly deviate from the experimental conditions. This standard deviation reference value is stored persistently in the device.

This approach is intended to be used for every sensor that requires calibration time. The pH sensor was mentioned as an example because it is a crucial sensor in our project.

3.2.5 Ph Sensor

To deploy the **pH sensor**, it was first necessary to make experimental readings using this device. The ESP32-S2 includes ADC (Analogue to Digital Converter) channels to make analogue readings. This input value, the ADC, is a value with 13 bits range, which means is able to read values between 0 and 8191. Zero means null voltage, while 8191 is the maximum voltage the MCU is able to read.

The ADC units in the ESP32-S2 support configurable attenuation settings to handle different voltage levels. When higher attenuation is set, the ADC's input voltage range is increased, allowing measurement of larger input voltages.

According to the documentation, 11dB (max) allows the chip to read from 0 - 2500mV [28]. This means, that with an ADC value of 13 bits, and an attenuation of 11dB, the ADC is almost always sensible to voltage change between 0 and 2.5V, excluding extremes as with voltages near 0V and 2.5V. The read values should never achieve reach these extremes, and so, it imposes no problem to the readings.

Since the pH sensor outputs 5V (maximum) we used a voltage divider to downgrade the analogue input voltage on the chip, to 2500mV (the maximum voltage that can be read by the MCU).

The formula for calculating the output voltage is:

$$V_{out} = V_{in} * \frac{R2}{R1 + R2} \quad (3.1)$$

To achieve a maximum output voltage (V_{out}) of 2.5V while operating with a maximum input voltage (V_{in}) of 5V, we carefully selected two resistances with equal power ratings. Our selection was guided by the consideration of minimizing current flow to conserve system energy, while simultaneously avoiding excessively high resistance values that could introduce signal interference, since the weaker signal can be easily

overwhelmed or distorted by the interfering signals, leading to inaccuracies or disturbances in the system's operation.

We opted for 47kOhm for the value of both resistor(R1 and R2). R1 is the resistor connected to the input voltage (V_{in}), and R2 is the resistor connected to the ground or reference point as shown in Figure 3.6.

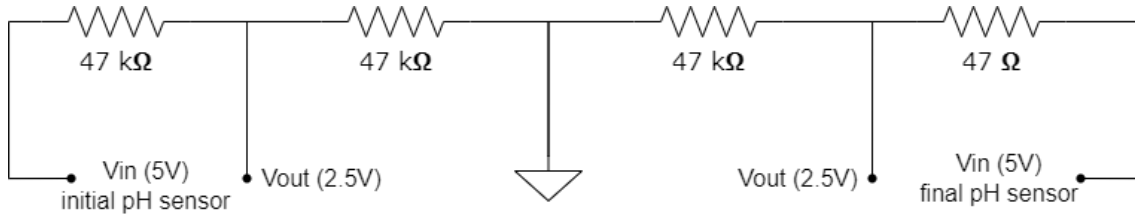


Figure 3.6: Voltage divider circuit used to achieve lower output voltage

To interpret the ADC values obtained from the readings as pH values, we had the option of utilizing one of two different approaches to establish this correlation.

The traditional one, which consists in the analysis of the characteristic curve of the sensor, which establishes the relationship between voltage and pH. It would also be necessary to align the voltage measurements from the MCU with the reference values obtained from the external device. This calibration process would involve the analysis of the data present in the Figure 3.7, to adjust the voltage read from the MCU to the voltage read from the oscilloscope.

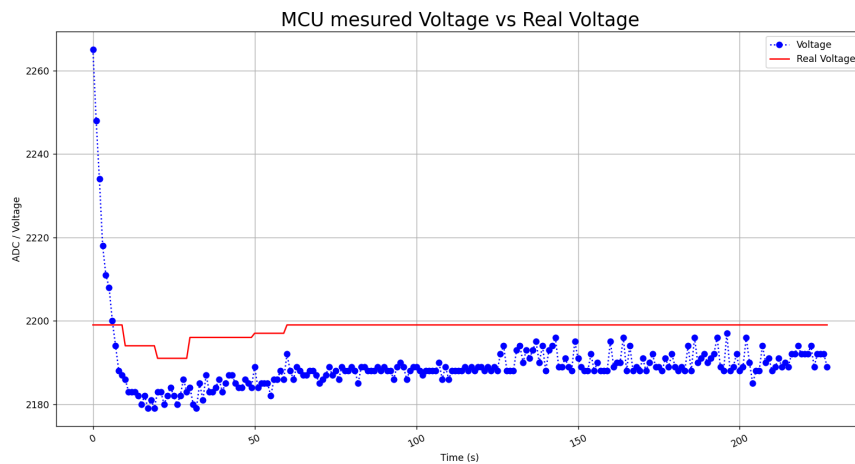


Figure 3.7: Graph illustrating the measured voltage with the MCU against the real voltage measured with an oscilloscope, of a liquid solution, pH equal to 8.

However, as we have liquids with known pH, we go directly from the ADC value to the pH value. According to the figure 3.8 the conversion between the ADV values and the voltage is almost linear, excluding both extremes. The voltage to pH transformation function is also linear.

PH values are relatively linear over a certain range (between pH 2 to pH 10), we need two calibration points to determine the linear line, and then derives the slope of the line so that we can calculate any pH value with a given voltage output, and therefore, ADC value.

Knowing this, we experimented a solution liquid, measuring pH equal to 4 and

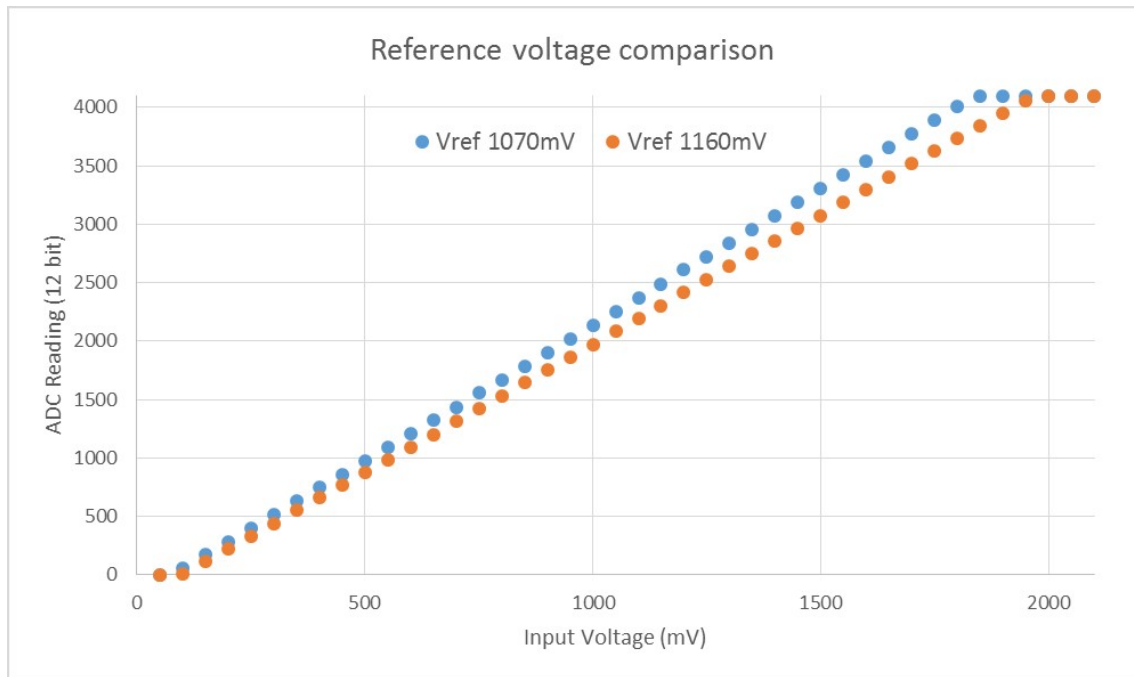


Figure 3.8: Graph illustrating effect of differing reference voltages on the ADC voltage curve.

read the correspondent ADC value in the MCU. The same was done with another solution measuring pH equal to 8.8.

PH deduction:

7955 ADC corresponds to pH 4.0

7184 ADC corresponds to pH 8.8

$$m = \frac{8.8 - 4.0}{7184 - 7955}$$

$$4.0 = 7955 \cdot (-0.00622) + b$$

$$b = 4.0 - 7955 \cdot (-0.00622) = \pm 53.4801$$

$$\text{ph} = \text{ADC} \cdot m + b$$

$$8.8 = 7184 \cdot (-0.00622) + 53.4801$$

We have a high level of confidence in the accuracy of the pH values of the liquids used in our experiments, since we validated them through the use of a commercial calibrated pH meter.

With this, it is possible to trace the slope, allowing the conversion from ADC to pH. Even so, after powering up the pH sensor, it is necessary to wait for the sensor to stabilize, as the figure 3.9 shows. Starting from the data point at the 102 second(ADC: 7947), we observe a sequence of consecutive data points where the ADC values remain within the range of 7947 to 7953 (± 3 units). All the values after 102 seconds consistently fall within the defined stabilization criteria of ± 3 units. Therefore, it was decided that after 102 seconds, the reading is considered valid.

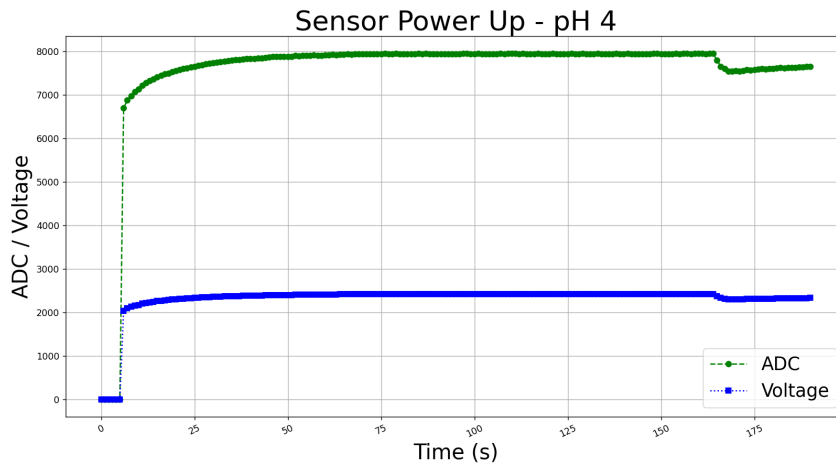


Figure 3.9: Graph illustrating the measured ADC and voltage of a liquid solution, pH equal to 4.

Ambient humidity and temperature sensor

To make the ambient humidity and temperature readings, the MCU component that handles this needs to respect the behaviour of the chosen sensor module, as described in the subsection 2.1.1.

During the development process, we conducted some research to identify existing implementations that could be leveraged, thus avoiding the need to write low-level code from scratch.

However, to determine if calibration was necessary, we conducted an experiment to verify if the values were close to the real environment condition. The test results indicated that after the 30 seconds to that took the sensor to stabilize, the values were correspondent to the room temperature and humidity, in Figure 3.10, the data relative to the reading of the sensor is presented graphically.

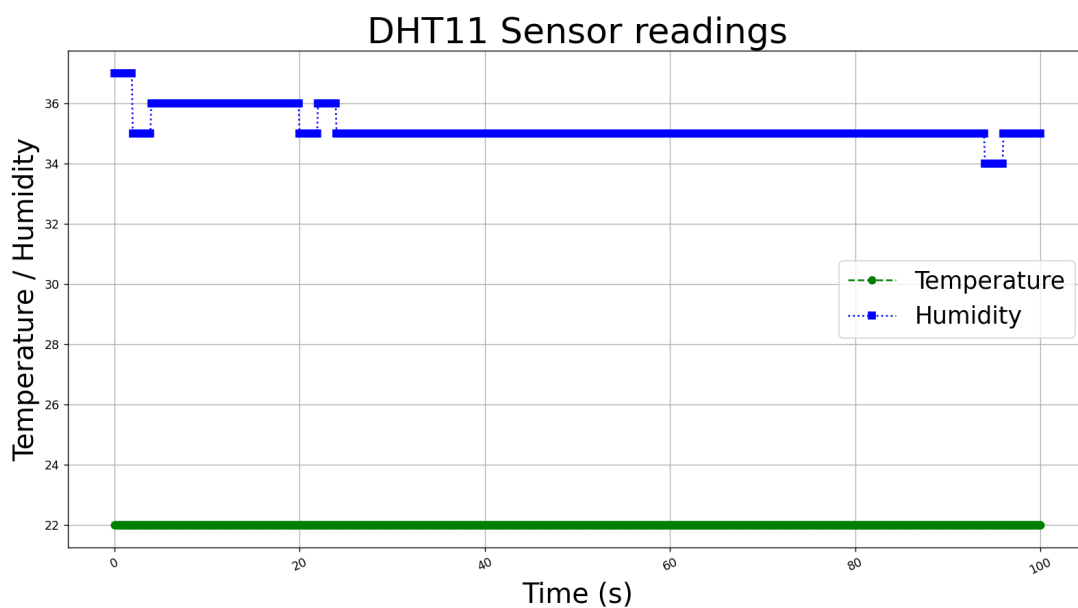


Figure 3.10: Graph illustrating the measured humidity and temperature, from with the DHT11 sensor, at room temperature.

Water leaks sensor

Relative to the water leaks sensor presented in the subchapter 2.1.1, the sensor will output a certain voltage, each time it is in contact with water.

At first, we decided to use the ESP32-S2 functionality of waking up the micro-controller whenever an input pin receives the digital 1 value from the sensor. We identified two problems with this approach:

1. After conducting experimental tests, it was concluded that the module, even when fully immersed in water, only outputs 1.7V. This voltage level is insufficient for the MCU to recognize it as a digital logic high. The graph in Figure 3.11 illustrates the measured ADC and voltage when the sensor is immersed in water and taken out;
2. The sensor would have to be constantly on, resulting in greater energy consumption. This goes against the established requirements, where the hardware is supposed to be energy-efficient. Additionally, it is assumed that any water leakage in the system would occur at a very slow rate.

Hence, we made the decision not to keep the water leakage sensor actively monitoring for any potential leakages. Instead, the sensor would operate similarly to the other sensors. When the MCU wakes up from its deep sleep mode, it will read the sensor value. If the sensor detects any water leakage, the event will be reported.

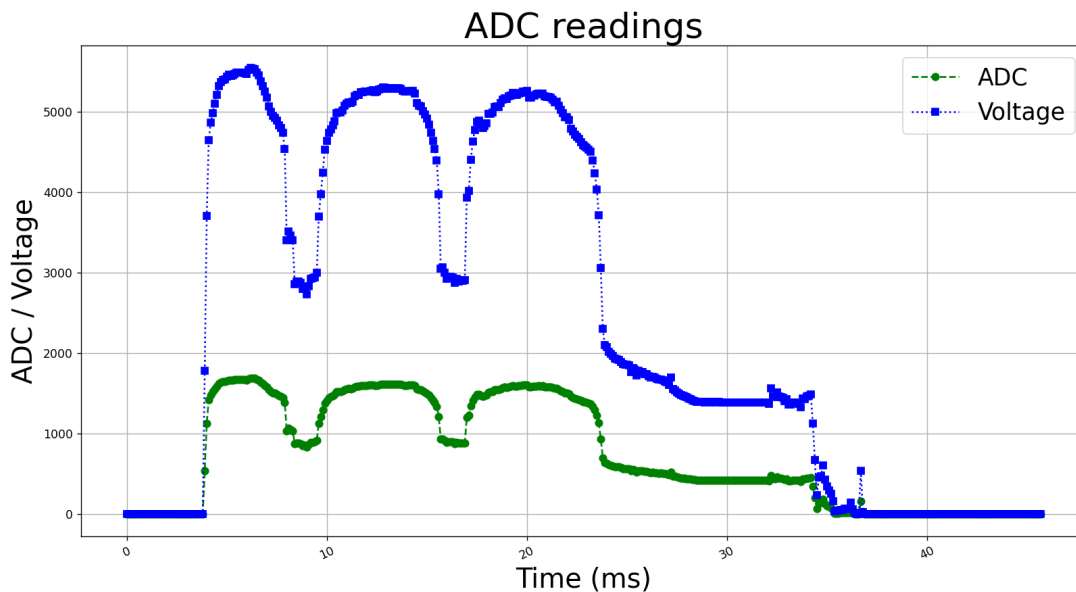


Figure 3.11: Graph illustrating the measured ADC and voltage when the sensor is immersed in water and taken out.

3.2.6 Assembly

All sensors, except the pH sensor, can be powered on with 3.3V to operate. Since each GPIO can be used as digital, operating on 3.3V, when outputting digital one, they can be used as digital gates, for each sensor. Therefore, each one is connected to an individual GPIO. Using different I/O pins to supply each sensor indepen-

dently offers two significant advantages. Firstly, it helps optimize power consumption by allowing each sensor to be powered on only when the chip requires their functionality. This reduces unnecessary power drain and contributes to overall energy efficiency. Secondly, this approach provides isolation in case of a circuit issue or malfunction in any specific sensor. If a problem arises in one sensor, only that particular sensor will be affected, and the rest of the system can continue operating without interruption. This isolation allows for easier troubleshooting and maintenance, as the problematic sensor can be identified and addressed without impacting the functionality of other sensors or components. For the pH sensors, which require a total of 5V to function, we utilized the only PIN capable of outputting 5V. This specific pin was employed to supply power to both pH sensors. The issue with this particular pin is that it cannot be controlled to reduce power through software. As a result, the pH sensors would remain operational even when the MCU was not actively utilizing them, leading to unnecessary power consumption.

To address this issue, we implemented a solution using a transistor to interrupt the power supply when needed. Since both pH sensors had similar power stabilization timings and were used almost simultaneously, we decided to utilize a single transistor for both sensors. A separate GPIO pin was assigned to control the gate of the transistor, allowing us to effectively manage the power supply to the sensors.

In terms of data connections, there is a total of 5 pins in use:

- The initial and final pH sensors connected to the pin 3 and 4, so it's possible to read independent values from each sensor. These pins have each one its own independent ADC channel, respectively, the ADC1_0 and ADC1_1 channels, allowing to read analog input voltage. Both pH sensors have their own resistances to lower down the output voltage, as explained previously;
- The water sensor, connected to the pin 3, also used as an analog channel - ADC1_2;
- The DHT11, serving as the temperature and humidity sensor, which is digital, is connected to the pin 11 - GPIO 9, transmitting data in a series of bits.

The pin 10 - GPIO 8 - is being used to serve as an external wake up, when the ESP is sleeping. When the pin receives the digital value 1, it triggers a reset in the ESP configurations, when he wakes up. This pin is only used when the MCU is sleeping. Figure 3.12 depicts the schematics of our assembly.

3.2.7 Wi-Fi

To connect the MCU to the Wi-Fi, two properties are required: **SSID** and **Password**. The SSID and the password are the credentials used to connect to the Wi-Fi network. Firstly, the MCU tries to recover the Wi-Fi configuration and, if available, tries to connect to the Wi-Fi. If not, the MCU will initiate the *ESP Touch protocol* to allow users to utilize the *ESP Touch* Android app to transmit the Wi-Fi configuration to the MCU.

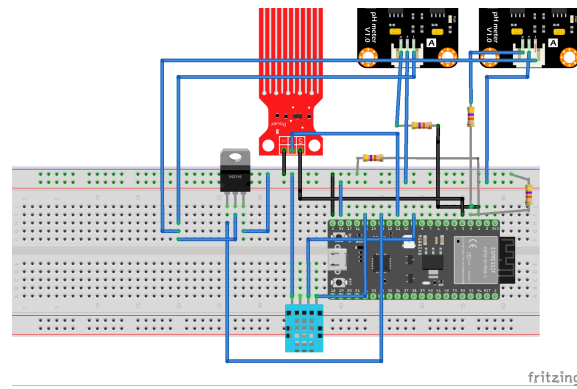


Figure 3.12: Current IoT Node assembly in the breadboard

Once the user sends the credentials to the ESP32-S2, the data will be stored in the persistent memory, for later to reconnect to the same SSID. Through the Touch Android application, the user has also the ability to send a custome string, as described in the subsection 2.1.4, which in our project we define this string as the unique **device ID** that corresponds to an identifier exclusive for each microcontroller.

To incorporate the ESP Touch Protocol and Wi-Fi modules into our project, we referred to the **SmartDoorLock-IoT** project mentioned in section 2.2. This project had already implemented these modules, so we utilized and customized them to suit our specific use case.

3.2.8 MQTT

To send the sensor collected data to the backend broker, we developed the MQTT'util module. Most of the implementation is taken from the extensive ESP IDF documentation, which deals with the MQTT connection to the target endpoint, and other low level details. Yet, we adapted with other procedures to send sensor data, according to our domain.

In our IoT system, where devices may operate in an unreliable network environment and energy efficiency is crucial, we have considered various MQTT data formats. While a compact payload format can contribute to system efficiency, we have opted to use the JSON format. Despite JSON potentially having a slightly larger payload compared to other compact formats, such as binary or compressed formats, we have determined that the benefits of its flexibility and ease of parsing outweigh the minimal size difference.

The content of the JSON file includes the device ID, the timestamp, plus the sensor value or the device wake up reason. If, for instance, the payload included a list of undefined objects, a different serialization mechanism would be required. Also, the MCU doesn't receive messages from the exterior. One inefficiency related to the JSON format is the de-serialisation process, which never happens in the current system.

One deficiency in our system is the need to configure the MQTT broker URL in the firmware itself. One possibility to solve this problem is to provide it through the ESP Touch protocol, using the already described Android App. This would be

necessary only for the first time the user interacts with the IoT MCU device.

3.2.9 Storage

To persist data in the ESP32-S2, the Non-volatile Storage ESP IDF library was used in this project. Non-volatile storage (NVS) library is designed to store key-value pairs in flash. It is a type of non-volatile storage device, meaning that even if it has no power, it will retain the information saved in it [29]. This way, even when the device is powered down, data won't be lost. In the MCU firmware, there is a module called `nvs_utils` to interact with the NVS partition, providing an interface to store and retrieve the device ID and Wi-Fi credentials.

3.2.10 Time

The ESP, sometimes, will have to obtain the current time, for instance when reading data from sensors.

There are two sources of time, in the ESP32-S2:

- The RTC Timer, which is capable of persist time over deep sleeps, but is not so accurate;
- High-resolution timer which doesn't persist time between deep sleeps but has a greater accuracy.

Both sources lose track of time when the device is powered off. Therefore, an external synchronization mechanism is required every time the MCU is powered on. This synchronization process is unavoidable and requires internet connection.

We opted for the RTC Timer since time accuracy is not important in this system. Moreover, the power consumption is greatly affected by WiFi transmissions, and thus, requiring external time synchronization all over again is unnecessary, given our requirements.

For time synchronization, we have selected the Network Time Protocol (NTP) to acquire accurate timestamps. NTP is a protocol used for synchronizing the time on computer systems over a network. It is designed to achieve high accuracy and reliability in time synchronization. NTP uses a hierarchical system of servers and algorithms to distribute time information and minimize clock discrepancies between devices.

Overall, the ESP will use the NTP protocol to sync the time, one and only once, after a power on reset, persisting time for the rest of its execution stage.

3.3 Backend

This component is shared by all the clients, and serves an IoT platform, supporting all IoT nodes scattered across the multiple clients using this system. It should be deployed in the cloud, so all IoT nodes have access to it. It is composed by the application server, the broker and databases.

3.3.1 Broker

The broker is used as an intermediary in the communication between all existent IoT nodes and the backend server. There are three topics used for communication:

- `sensor_record` to send a single sensor recording.

```
{
  "device_id": "AdPGFVs",
  "value": 8.0,
  "timestamp": 1656330216,
  "sensor_type": "temperature"
}
```

- `device_wake_up_log` used by the MCU to report the reason for wake-up. Common values are timeout reason, exceptions/errors, power-on, software or hardware cause, etc.

```
{
  "device_id": "AdPGFVs",
  "timestamp": 1686330216,
  "reason": "power on"
}
```

- `error_reading_sensor` used by the MCU to alert if there was a problem while reading any sensor.

```
{
  "device_id": "AdPGFVs",
  "timestamp": 1686330266,
  "sensor_type": "initial -ph"
}
```

3.3.2 Application Server

The server computes all system involved data. All the business logic is enforced here. This application serves two purposes:

1. Exposes a Web API, allowing client applications to interact with the system;
2. Store and compute all system data, enforcing business logic;
3. Receive, analyse and store the data that coming from the Broker module.

It also exposes a Web API, so that Frontend applications can interact with the system.

The Figure 3.13 shows the main modules involved:

1. Authentication and authorization module will be used when the request is attempting to access protected resources;
2. Other interceptors, such as request logging interceptors;
3. The controllers to define the API operations;

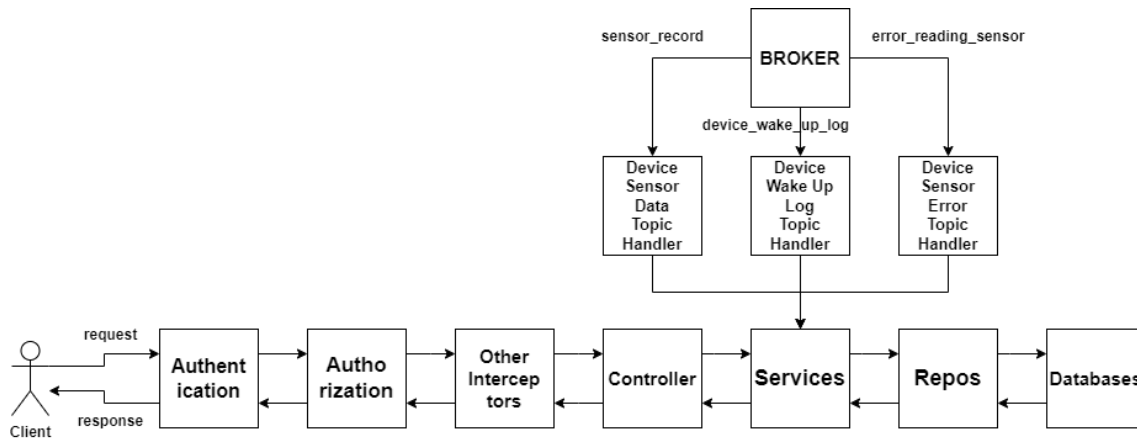


Figure 3.13: Spring server overview

4. The services display available system actions which will guarantee atomicity and enforce business logic while accessing and manipulating system data;
5. The repositories implement the database "drivers", meaning they handle all the logic associated with the database access;
6. The MQTT topic handlers, responsible to receive with data coming from the broker, and deal with them. This includes analysing sensor data, storing coming records, etc.

Each time a client sends a request, the controllers will use the services to fulfil the request. Likewise, when the services need to access or store data, they will use the repositories to interact with the physical databases. In essence, the request will follow a left-to-right flow, while the response will flow in the opposite direction.

3.3.3 System Logic

Users

As with any other system, every user needs to have a unique identifier. We opted to use a UUID (Universally Unique IDentifier), a unique identifier with 128-bit, since the probability of replications is very small.

We also opted to define two types of users, each one with different capabilities:

- The "admin" role in the system has privileged access to all users and devices, including their sensor data and logs. However, this user is not authorized to create devices. The primary objective of the admin is to monitor other users and their corresponding devices. This role is typically assigned to a member of the company responsible for manufacturing the pH neutralization mechanism. Their purpose is to gather data from all devices to improve the product and provide valuable insights. This user cannot be created using the Server API. Its creation is automated by the backend Server when this is initialized;
- The client, which is able to create devices and is authorized to manage his own devices. This user can be created using the backend Server API.

The "admin" role was not a request by the company, but an implementation approach that we decided to take. In addition to automating the monitoring of the neutralization filter, this project aims to extract relevant data for evaluating the system's efficiency. The admin role, plays a crucial role in this aspect by having access to all the data collected from the system. This access allows the admin to analyse and evaluate the performance of the system, enabling improvements and optimizations based on the extracted insights.

Devices

Each device needs its own ID, and a 32-bit ID (2^{32} possibilities) is more than enough for our use case. This extensive range of identifiers ensures that each user can be uniquely distinguished within the system, providing a high level of identification accuracy and minimizing the chances of duplicate identifiers.

This system is designed to allow clients to configure their own devices. During the configuration process, the device identifier generated by the web app needs to be manually passed to the device using the Esp Touch Android app. Since the client will have to manually write the device ID using his private phone, a giant ID wouldn't be suitable. In order to minimize the effort required, we generated a unique identifier consisting of 6 letters, from A to Z, lower and upper case, which is the minimum required length to represent all the 2^{32} possibilities imposed above. This results in a total of 4,294,967,296 possible combinations.

Each device is associated with an alert email, defined upon creation. This will be used to notify the user if the filter may need human intervention.

3.3.4 Sensor Analysis

Upon receiving sensor data from the broker, the server is responsible for analysing the data. The analysis involves evaluating the received sensor records to determine if any maintenance is required for the filter mechanism. There are many possibilities to achieve this. The Server only needs to know the thresholds (boundaries) associated with each sensor type. For example, the pH lower and upper boundaries might be 6 and 7. One possibility would be for each device to send, along with each value record, the sensor thresholds. The server would receive the value and the respective thresholds and process it immediately. The main disadvantage is that it requires more bandwidth, which deeply impacts the power consumption on the MCUs. The advantage would be, it would be possible to customize threshold values among different devices. This system assumes that all filters belong to the same company, and thus, they should function under the same conditions. Therefore, the current solution is to store the sensor thresholds inside a configuration file, in the Backend server. This implies that all devices sensor data will be treated similarly.

3.3.5 Spring Configuration

One notable feature of the Spring framework is its robust dependency injection mechanism. This feature offers a powerful abstraction for programmers when implementing software layers within the Spring framework.

This system effectively utilizes this feature. For instance, the server repository classes rely on a *Handle* object, which represents a connection to the database system and is closely linked to the database. During application testing, such as integration tests, it becomes necessary to provide a different database to the *Repository* class. This is where the Spring configuration comes into play. In this project, various configurations are defined to accommodate different databases, depending on whether the application is running in a production or testing environment.

3.3.6 Authentication and Authorization

Most Spring Server API operations request credentials to access resources. Therefore, the Server API allows the client to register itself. To identify logged in clients, a cookie must be present in the http request header, containing the session token.

There are two options to register a new user:

- Use Google authentication. The server uses the OAuth2.0 protocol to allow third client applications to authenticate in the Server API. If the user is using google authentication for the first time, the Server will create a new user, based on the Google profile. In this case, no password will be stored since the Google will take care of the authentication process. This operation will also place a session token in the cookie;
- use standard user creation, which requires an email address and user password. The user must log in, afterwards, to access protected resources.

In order to control requests coming to the server, a Spring interceptor is used to perform authentications and authorization against the request.

The interceptor analyses the request containing the cookie with the access token and comparing against the database. When not valid, the interceptor determines that the request should be terminated because the user is classified unauthorized, responding with a *401 Unauthorized status code*. This means that the user's request is immediately halted, and a response indicating the lack of authentication or authorization is sent back to the client.

The 401 Unauthorized status code is a standard HTTP response status code used to indicate that the client request has not been completed because it lacks valid authentication credentials for the requested resource [30]. By terminating the request and returning a 401 Unauthorized response, the client is notified that they need to provide valid authentication credentials to access the requested resource. This approach ensures that unauthorized access attempts are promptly rejected, and the client is informed of the authentication requirement. It helps enforce security measures and protects sensitive resources from unauthorized access attempts.

To enforce authorization, the same interceptor tries to authenticate the user and evaluate if the user possesses the necessary role defined in the annotation to be allowed to proceed with the request. If not allowed, a 403 Forbidden status code is returned. The 403 Forbidden status code is an HTTP response status code that signifies that the server understood the request but refuses to authorize it. It is used to indicate that the server has explicitly forbidden the requested action, even if the user is authenticated.

3.3.7 Controllers

In Spring, a Controller class is always proceeded with a *@Controller* annotation. In particular, in our application, the *@RestController* annotation is used over the other one, because it pre-defines a set of standard assumptions, common to Rest API implementations. For example, the *@RequestMapping* methods assume *@ResponseBody* semantics by default.

For organization reasons, the controllers are separated into multiple components, described in the architecture.

As an example, the module *DeviceController* includes the method *addDevice*, which has the *@PostMapping(URI)* annotation. By including the *Authorization* annotation, the annotated method becomes the designated handler that will be invoked when a client sends a request to the specified URI (Resource Identifier) mentioned in the annotation.

The method takes a single parameter of type *User*. This informs Spring to resolve the requested parameter automatically before executing the handler. By requesting this parameter in the method signature, the *Authentication Interceptor* described in the previous section ensures that only authenticated requests reach the controller.

All handler methods have another type of annotations: API documentation annotations. These are used to generate automatically the Swagger specification API documentation[31].

Inside the controllers, an *Authorization* annotation is used. This is not part of Spring. This one is an application domain annotation, and is used to mark the controller methods that require the user to have a specific Role: User or Admin.

3.3.8 Database

Due to the nature of this project, there are two types of data to be persisted:

- Aggregate data: Users, devices, etc;
- Time based data: Sensor data

A relational database (RDB) was chosen for storing data that involves relationships, such as users having devices and devices having logs. A relational database is specifically designed to enforce structure and maintain relationships between data

entities. Additionally, when dealing with sensitive data like user information, it is crucial to handle it with care.

In this project, *PostgreSQL* was selected as the relational database. It is an open-source, powerful, and highly reliable database system. Its robust features, scalability, and strong community support make it a suitable choice for handling complex data structures and ensuring data integrity [32].

On the other hand, we chose to use the *InfluxDB* database to store sensor record data. This is a time series database, which is specially designed for this kind of data, where each recording involves a value and a timestamp.

To construct each database, our approach began with the design of their data models. These models serve as a structure with the purpose of organizing data elements and establishing standardized relationships among them. By doing this, we are selecting the requirements for our specific domain.

The Figure 3.14 represents the schema of a relational database designed to manage user, device, log, sensor error, token, and password information. The schema defines the relationships between these entities, ensuring the proper association and integrity of data within the system.

Each user can have multiple devices in operation. It has been determined that each device must be exclusively associated with the email address of the supervisor. This email address is provided during the initial registration of the device and serves as the designated recipient for all alerts related to that particular device. Each device can have a set of logs and sensor errors. This can only exist when associated to the device. The same goes for the tokens and passwords, as they can only exist tied to a user.

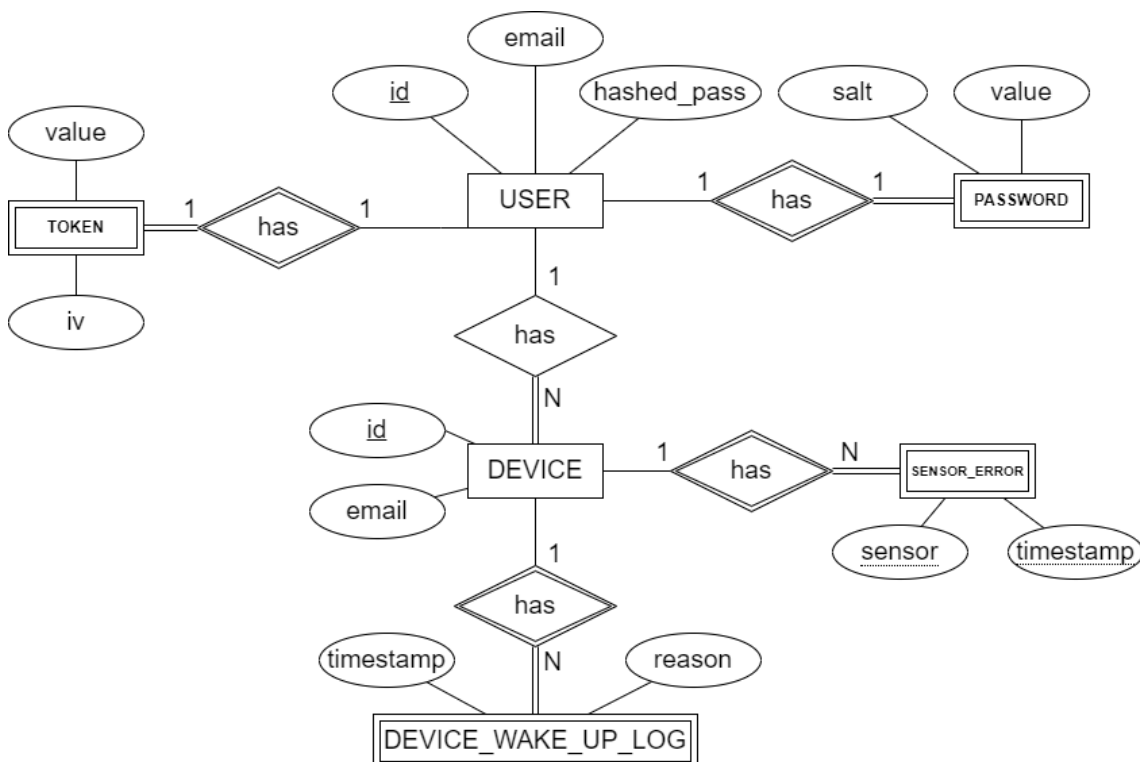


Figure 3.14: Postgres DB Entity Model

To summarize the relationships:

User and Device: One-to-Many (One user can have multiple devices). Device and Logs/Sensor Errors: One-to-Many (One device can have multiple logs and sensor errors). User and Tokens/Passwords: One-to-One (One user has one set of tokens and one password).

These relationships ensure the proper association and integrity of data within the system, allowing for effective management and organization of user, device, log, sensor error, token, and password information.

As InfluxDB is a non-relational database, the standard relational data representation approach does not apply. However, the following diagram offers a clear understanding of the type of data it stores:

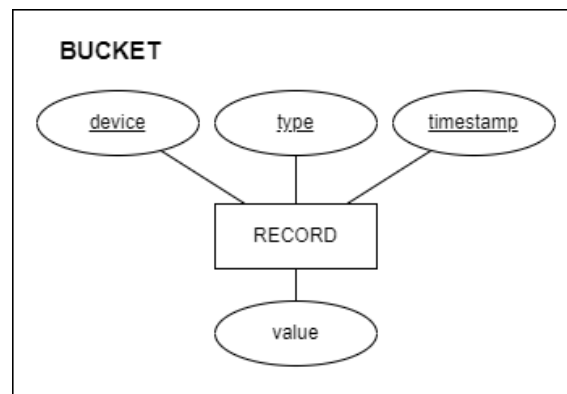


Figure 3.15: Time Series DB representation

A record always consists of the device, the type or sensor, the timestamp, and the corresponding value or reading.

3.4 Website

As the third system component, the Web application appears as a user-friendly entry point to the system. It interacts with the exposed Web API, in the Backend, to manipulate system data.

The main website architecture is represented in Figure 3.16, it presents an overview of the different modules and connections that constitute the website.

The App module is responsible for loading the Router, which is responsible for selecting the component that the user will see.

The majority of components within the application rely on the services layer to access the backend server. This allows them to request and manipulate system data.

The services layer acts as an intermediary between the components and the backend server, handling the communication and data operations on behalf of the components.

The main objective of the website is to provide an interactive user interface to allow the buyers of the neutralization mechanism a way of graphically visualizing the data that is being collected.

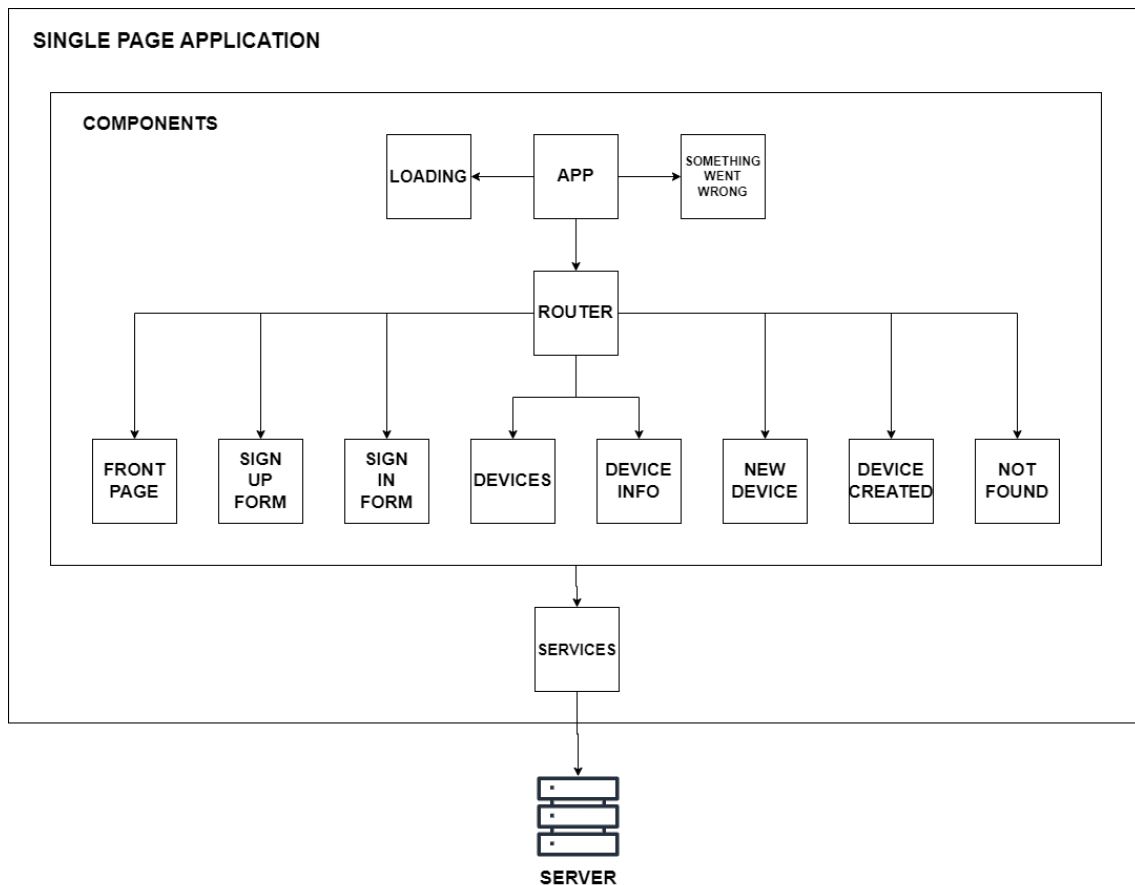


Figure 3.16: Single Page Application components diagram

3.4.1 Website Behaviour

When the *React App* component is loaded, the first thing to do is to load all the siren API information, from the Backend.

In a Siren-based API, resources are represented as JSON objects with specific properties such as "class," "properties," "links," and "actions." The "links" property provides links to related resources, allowing clients to navigate between them. The "actions" property describes the available actions or operations that can be performed on a resource, along with their associated input parameters. This is used to facilitate the Frontend development, since it is useful for the website to have to know only one the base URL and from this endpoint, it should be able to get all the API information, in order to navigate all the API endpoints, without having prior knowledge of the API structure. If this information cannot be obtained, the website will only show a static error message. This behaviour is because if the SPA is not able to obtain all the expected Siren Info, it would cause a bad user experience. For instance, it might not be able to access devices from the Backend. Therefore, the action is to display a message to the user, showing that something went wrong. A future idea would be to register the SPA developer to receive alerts each time this happens, since is considered a severe error.

One possibility which wasn't yet implemented is the use of local cache to store the API information. This ends up being a trade off. In a way, it is possible to make refresh loads without erasing local information, such as the siren information, or the

login state. On the other way, it might be a risk, since the cache could become out-dated. This is very unlikely in the case of the API/Siren information. But could emerge as an issue on the latter case (login state).

If the Siren information is obtained successfully, the SPA will set up the React Router, and the authentication state will be initialized.

3.4.2 Protected Resources

Most website resources/pages are protected, meaning the standard users should not be able to access them, but only authorized ones. Therefore, the website provides a sign-up/sign-in page, for authentication purposes. After users login, they will be redirected to an authorized page. Yet, nothing forbids the user from trying to access a protected resource directly, for instance by inserting the page URL manually in the browser search bar (deep linking). For that reason, there is a special component - **RequireAuthn** - that wraps the protected component, ensuring that the user is logged-in. His behaviour is demonstrated in the ASM chart represented in Figure 3.17. By using this component, it ensures protected components are only accessed by authenticated users.

The login state is managed by the *AuthnContainer* component, and accessible to most SPA React components, so they can check if the user is logged in, and use that information, for example to make an API request.

The reader may be thinking: "the authentication state might be safe in the *AuthnContainer* component, once the user logs-in. But what happens if the user refreshes the page, and the authentication state is lost?". Should the user have to log in again?

In our system, we store a token inside a cookie to hold the login state. This is used, not for the browser to know that it is logged in, but for the server to know the user is logged-in. The source of truth lies in the token validity. Since the cookie, sent by the server, cannot be accessed in JavaScript, to protect from cross-site scripting (XSS) attacks, the browser/SPA needs to make an API request to know if the client is authenticated or not. This happens when the user uses deep linking to access a protected resource, and the authentication state is not yet known by the SPA. Yet, the token stored in the browser might still be valid, and he should not have to go through all authentication steps again. Similarly, to make the logout, the client browser makes an API request, so the server can remove the cookie, therefore eliminating the client session.

3.4.3 Error Handling

Most unpredictable events will come from the Backend server (ex: when the user requests its own devices, sensor data, etc). In these situations, after making a request, the response could be a 400s or a 500s. The response will, most times, be a problem+json, and have a detailed description about what caused the error outcome. Despite this, it is not always practical for an SPA to know all possible errors, and know how to deal with them in particular. Some errors are expected. For instance, the React component that shows the login/sign-in page and makes the Backend request to

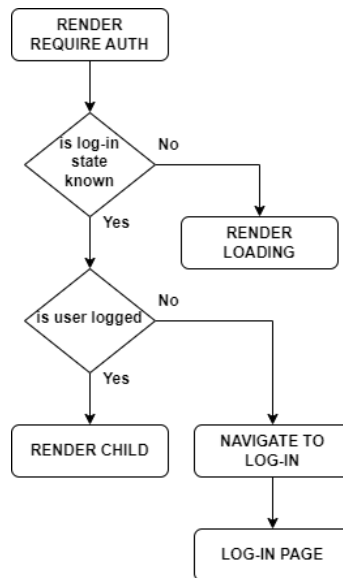


Figure 3.17: Authentication control in the SPA

create a User, will expect possible errors related to the credentials. Yet, in other request types, sometimes the error is a 500 internal server error and the website might not know how to deal with it. Therefore, the solution was to use a React component - *ErrorContainer* - that wraps another (the component that the user should see if everything goes well) and re-renders itself each time an error happens. By doing this, it is possible to control the user experience, specially when something unexpected happens, by not just letting the error pass unnoticed.

3.4.4 Services

To facilitate website development, we created a Service interface along with two implementing classes: *FakeServices* and *RealServices*. The *Service* interface defines the contract for interacting with the backend server API. The *FakeServices* class was specifically designed to mock the services provided by the backend server API to help with the frontend development.

However, in this report, we will not go into detail of the *FakeServices* class since its primary purpose is to provide a mock implementation for frontend development. Instead, we will focus on the *RealServices* class, which represents the actual implementation that communicates with the Backend Server API.

The majority of the service functions need to make an API request, hence they need to know the specific URL, the HTTP method and other information. The backend server API works with the media type: Siren. In particular, the only known URI endpoint - *siren-info* - responds with all the information the website needs to know to navigate the API. The function *getBackendApiInfo()* will make a request to this endpoint, and extract all necessary information. This includes the Siren Links and Actions, allowing the website to function.

When a service function needs a link or action that is not yet known, it throws an exception, so the caller can know the problem. They can also throw exceptions for other reasons. For instance, the *fetch* function will throw an exception if the API

response is not the one expected, i.e. 400s. If the response is a 400s, the fetch will try to convert the error response into a *Problem JSON*, throwing an exception with the *Problem* title, allowing the caller to know more information about the request error response.

3.5 System Tests

3.5.1 Backend

The central component in the Backend is the Spring Server, and so, extensive tests were implemented to access its well execution.

In our development process, we have placed a strong emphasis on automated testing, particularly for the server-side layers of our application. This includes testing the Services, Repositories, Domain, and Controllers through integration testing.

3.5.2 MCU Firmware

To test the MCU firmware, we opted for simple programs that need to be loaded in the hardware. These are not automatic tests, unlike in the Spring Server. For instance, to calibrate the pH sensor, a simple program that only reads from the sensor, is loaded into the hardware and the developer observes the logs, which print the values. Next, the values are copied and pasted in a .csv and loaded by a python script to display the values, and thus, providing a more comfortable analysis.

Another example of testing is to use a python script that emulates the hardware device to send synthetic sensor data to the Backend application server, so it is possible to test the server receiving sensor data, without relying on the real hardware.

Power Consumption

As emphasized throughout this report, power consumption plays a crucial role in the design of this system. This holds true not only for IoT devices in general, but particularly for this specific project, as per the requirements outlined by the company.

During the development phase, utmost importance was placed on implementing various power-saving techniques. However, the effectiveness of these techniques could not be determined without conducting proper experimental evaluations. To address this, a device called the UM24C USB tester was utilized to measure the power consumption of the ESP32-S2 during different execution stages. This equipment boasts a current measurement resolution of 0.001A, which provided sufficient precision for conducting meaningful experiments. The collected values, as presented in the table 3.2, represent the instantaneous current readings obtained during the evaluations.

To evaluate the power consumption of our system, a dedicated program was developed to simulate various operational modes of the MCU. Each mode was tested individually to assess its power requirements. Upon starting the program, the specific running mode was initiated and monitored for a duration of 30 seconds. This

approach allowed us to accurately measure and analyse the current consumption of the system under different operating conditions.

Current measurement in different states	
	Consumption
Active mode idle	27ma
Power on DHT11	29mA
DHT11 readings (spaced by 1 second)	29mA
Power on water sensor	27mA
Read from water sensor (no water detected)	29mA - CALCULATE MAX
Deep Sleep	10mA
Wi-Fi and MQTT setup	0.61A
Wi-Fi reconnect timeout (max tries)	78mA
Active mode Wi-Fi idle	31mA
Active mode Wi-Fi sending MQTT	39mA

Table 3.2: Results for the different MCU consumption stages

The tests yield readings way above the ones indicated in the respective data sheet. This difference can be explained by the presence of LED's in the ESP32-S2 and associated sensors. These LEDs cannot be turned off. During development, it is convenient to have them turn on to indicate when the sensor is being used, for debug purposes. Yet, from the experimentations made, it is noticeable clear the existent differences between different MCU profiles: active, with and without Wi-Fi, deep sleep, etc.

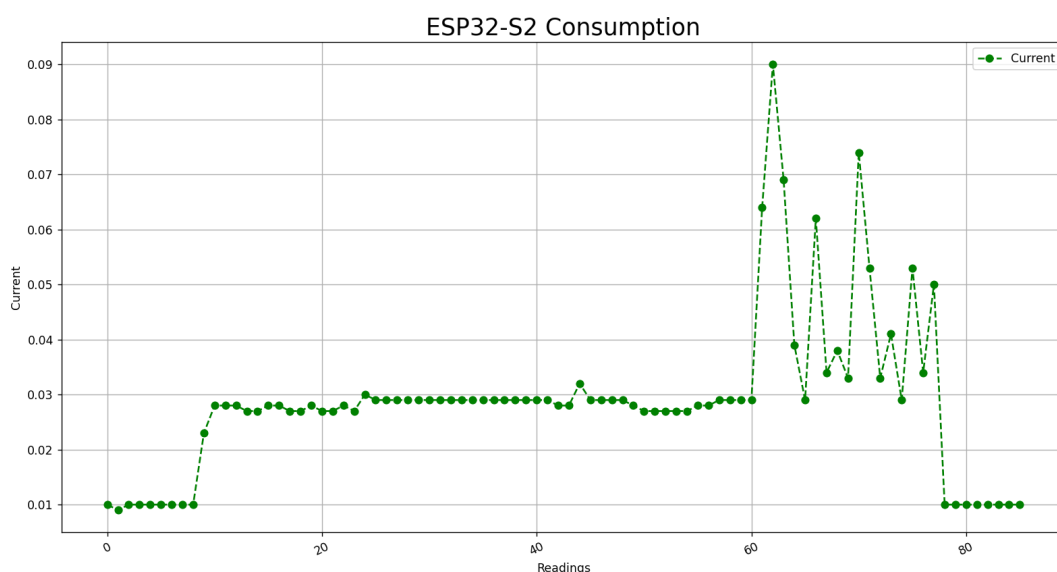


Figure 3.18: Graph illustrating the ESP32-S2 energy consumption measured when the MCU wakes up to read from all sensors

As explained in the subsection 3.2.1, the MCU will wake up and do one of two

reasons:

- to read from all sensors. This iteration will always use Wi-Fi to make data transmissions;
- to only check for water leaks. This iteration will only make transmissions if water leakage is detected.

The image 3.18 shows the instantaneous current consumed being consumed. In the initial and final stages, the MCU is in a sleep mode, resulting in relatively low power consumption. Upon waking up, the MCU engages the sensors to perform readings, leading to a moderate level of power consumption. However, when the Wi-Fi functionality is activated for data transmission, the power consumption significantly increases, indicating that this stage predominantly influences battery usage.

Power Range

To estimate how much the device battery is going to last, a test was conducted, in which the normal functioning of the MCU (excluding sleep modes) was left running for three straight hours. The test procedure involved the MCU operating in a continuous loop. In each iteration, the MCU alternated between two actions: reading from the water sensor only, and reading from all sensors and performing a data transmission. Yet, to simplify this test, we will treat every iteration as the same. Some stages that only happens once (ex: sensor calibration and time synchronizing) were left off from this test. Also, no exception was drowned because that would change the normal execution.

The number of cycles (wake-ups) was registered, along with the total accumulated consumption.

Stage	Total Consumption	Total Iterations
ESP working (wake up)	80mAh	149
ESP sleeping	30mAh	-

Table 3.3: Results for the different MCU profiles

Based on the project requirements, the ESP32-S2 is programmed to wake up at a 6-hour interval. To accurately estimate the total power consumption, it is necessary to consider the energy usage during both active cycles and deep sleep periods. With a 3400mAh battery capacity, it is crucial to ensure that the total consumption does not exceed this limit.

The consumption calculation is as follows:

Total Consumption =

$(\text{Energy per Cycle} + \text{Energy per 6-hour Deep Sleep}) \times \text{Number of Wake-ups} = 3400, \text{mAh}$

Here, "Energy per Cycle" refers to the power consumed during one complete cycle of operations, including sensor readings and data transmissions. "Energy per 6-hour Deep Sleep" represents the power consumption during the deep sleep state

for a duration of 6 hours. "Number of Wake-ups" corresponds to the total number of wake-up events throughout the project's duration.

By solving this equation, we find that there are 53 wake-up events in total, which translates to a period of approximately 14 days. This calculation enables us to ensure that the battery capacity is sufficient to support the desired operation of the system within the specified timeframe.

3.6 Security Measures

Ensuring the security of the system is of paramount importance in any application, especially when dealing with sensitive data and user authentication. In this section, we will discuss the security measures implemented to protect the system from unauthorized access and potential vulnerabilities.

3.6.1 Password Storage

. In our system, we prioritize the security of user passwords by storing them as hashes instead of clear text. A hash is a computed value derived from the original password, which adds another layer of protection. By securely storing passwords as hashes, we can maintain a high level of confidence in protecting user data within our system. In the event of unauthorized access to our system, the use of password hashing ensures that passwords remain secure. Additionally, hashing passwords enhances user privacy, as none of the database administrators or personnel can view the passwords in clear text form. To further enhance password security, we have implemented a technique called "salting" in addition to password hashing. When computing the hash value of a password, a randomly generated value (known as a salt) is added to the password before hashing. This salt provides an extra layer of randomness to the resulting hash. As a result, even if two users have the same password, their hashed values stored in the database will be different due to the unique salts applied. This approach significantly strengthens the security of passwords and mitigates the risk of password-based attacks, such as *rainbow table attacks* or *pre-computed hash attacks*. In a rainbow table attack, a large precomputed table of hash values and their corresponding plaintext passwords is used. The attacker compares the hashed password from the target system's database with the entries in the rainbow table to find a match. Precomputed hash attacks, on the other hand, involve the creation of a precomputed table of hash values for a large set of possible passwords. This table is generated in advance, allowing the attacker to quickly look up the hash value in the table to find the corresponding password. By salting the passwords, we introduce a random and unique value for each password before hashing. This prevents the effectiveness of both rainbow table attacks and precomputed hash attacks. Even if an attacker has access to the hashed passwords in the database, they would need to generate separate rainbow tables or precomputed hash tables for each unique salt, making the attack significantly more difficult and time-consuming. We can see a visual representation of how salts act in the figure 3.19



				
Password	p4s5w3rdz	p4s5w3rdz	p4s5w3rdz	p4s5w3rdz
Salt	-	-	et52ed	ye5sf8
Hash	f4c31aa	f4c31aa	1vn49sa	z32i6t0

Figure 3.19: Visual representation of the function of the salt [1]

3.6.2 Website

When a user attempts to create an account on our website, they are presented with two options: signing up through the Google authentication option or using the built-in registration mechanism provided by our web application. If the user chooses to sign in via Google authentication, a separate page will be displayed, requesting their permission to access their data. If the user grants permission, our application will retrieve their Google email and store it in our system, creating a new account using that email. If the user chooses to create an account through the website mechanism, they will be prompted to provide an email and password. To ensure password security, the password must meet certain requirements to be considered valid, such as having a minimum length of 8 characters, containing at least one digit, and at least one uppercase letter. This approach is implemented to discourage the use of weak passwords, as research has shown that common or easily guessable passwords are more susceptible to being compromised [33]. To prevent the creation of fake or unauthorized accounts and to add an extra layer of identification verification it is required that the user must confirm their email address through a verification code sent to the email supplied by the user, this process helps ensure that the email address provided during registration or account creation is valid and belongs to the user. This confirmation process also provides a barrier to scripts or bots with the objective of creating multiple accounts on the system, since the confirmation code is sent to the user's email address, automated processes would struggle to complete the verification step.

To maintain user authentication in our application, a token is utilized, as explained in the previous subsection on Protected Resources. This token is securely stored in a cookie and serves as a unique identifier for the user. However, it is crucial to protect this token from potential attackers who may attempt to hijack it through Cross-Site Scripting (XSS) attacks. XSS attacks involve injecting malicious scripts into web pages to gain unauthorized access.

To mitigate this risk, we employ the HTTP-only attribute to safeguard our cook-

ies. By setting the HTTP-only attribute, we ensure that JavaScript running on the page cannot access or modify the value of the HTTP-only cookie. This fortifies the security of the token stored within the cookie, making it significantly more challenging for attackers to exploit XSS vulnerabilities and impersonate users.

3.7 Deploy

This system is composed of three main components, described in the System Overview. The Backend and Frontend component should be accessible to all the system users. Therefore, we chose to deploy the system on a Virtual Machine (VM) accessible from anywhere. Since there are no free services of this kind, a pay-to-use service had to be chosen. We opted for the GCP compute engine service, since a 50\$ coupon was provided to us, allowing us to allocate a VM with no further costs.

To provide isolation between the different system layers, we opted to use docker engine and build containers for each service:

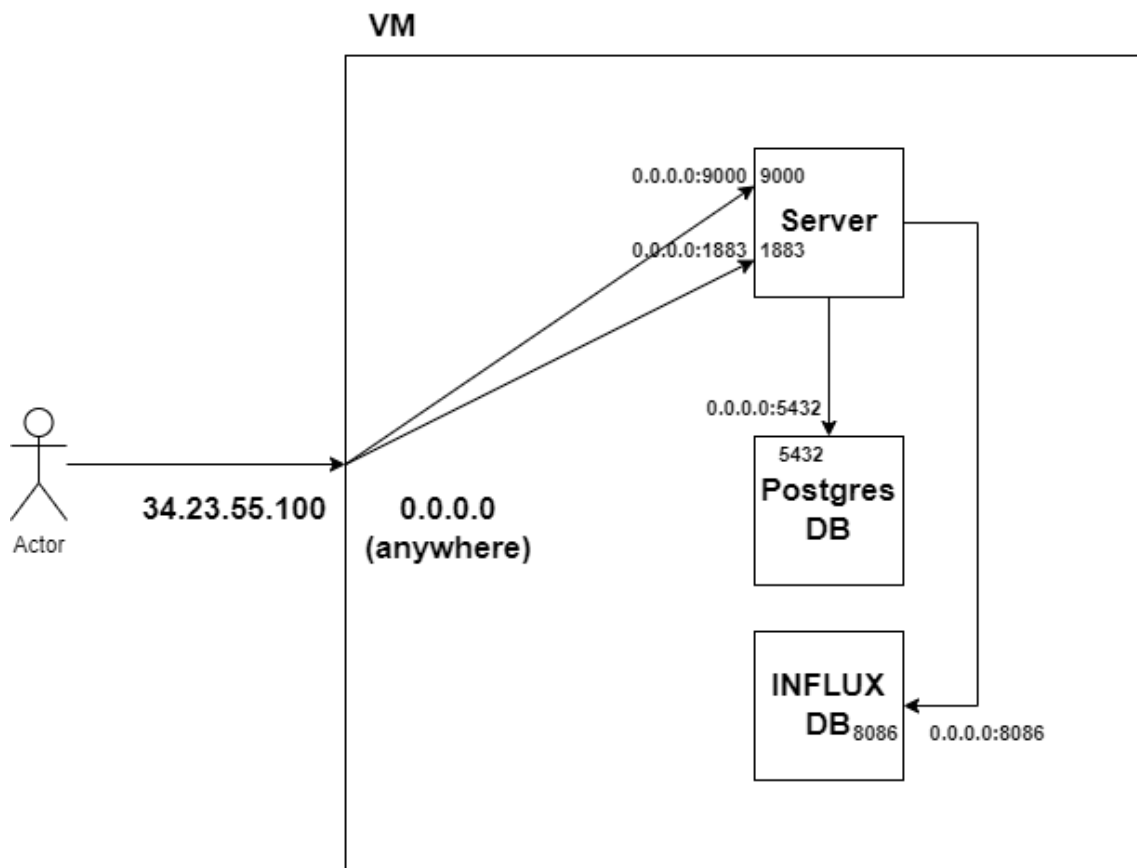


Figure 3.20: Containers for each system service layer

- The Postgres and Influx DB are two individual containers, built from an available open source image, supporting data storage;
- The Spring Server, receives the Backend requests and also serves the static website, and the JavaScript file. The Webpack was used just for development purposes. To make the deploy, Webpack was also used to package the several

JavaScript files into a single one, served here, in the Spring Server. The broker is currently being deployed from the Spring Server;

4

Conclusion

During the development of the project we had the opportunity to exercise the technologies we learned throughout the course such as React and Spring Framework and that are widely used in the industry.

We also had to learn to use other technologies and protocols such as time series database and MQTT protocol and beyond that we had contact with hardware, more specifically the programming of the ESP32-S2 microcontroller, which was one of the main challenges during the project since both elements of the group doesn't have any experience in this area.

Despite these challenges, the group managed to successfully meet the requirements that were proposed for the project.

4.1 Future Work

In this section, we discuss potential areas of improvement and future work for the system.

4.1.1 Hardware Adaptation

The current implementation utilizes the development kit of the esp32-s2, which includes LEDs that were useful during development but are not necessary for the final hardware system. As part of future work, the hardware design should be adapted to remove these LEDs and any other components that are not essential for the system's functionality. This will help optimize the system's form factor and reduce any interference caused by unnecessary components.

4.1.2 Dynamic Configuration

Currently, the URL of the broker, which serves as the connection point between the MCU and the server, is statically referenced in the code. To enhance flexibility and configurability, a future improvement would be to implement dynamic configuration for the broker URL.

4.1.3 pH Transformation with Temperature Compensation

The transformation of pH values is influenced by the temperature of the liquid being measured. In the current implementation, the system does not extrapolate or adjust the pH readings based on the temperature of the water. As a future enhancement, incorporating temperature compensation algorithms can provide more accurate and reliable pH measurements by compensating for the temperature effect. This would improve the overall precision and reliability of the system's pH measurements.

4.1.4 Hermetic Enclosure

To ensure the longevity and reliability of the hardware components, it is recommended to protect the device in a hermetic enclosure. This will safeguard the system from environmental factors such as dust, moisture, and physical damage. By creating a sealed enclosure, the hardware components will be shielded from external elements, ensuring the system's durability and longevity in various operating conditions.

4.1.5 Scalability

To deploy the MQTT broker, a dedicated container should have been used. This would allow us to scale up the broker independently from the server. Yet, at the current moment of writing, this wasn't done. The broker is currently being served from the Spring Server, and therefore, dependent from its availability. This is because, during development, it is easier to start the broker directly with the server, avoiding the need to launch it as another service.

4.1.6 Water Flow sensor

Despite the project requirements not requiring the existence of a water flow sensor, we wanted to integrate it due to the possibility of exploring more how the neutralization system is behaving itself. By collecting this kind of data, it becomes possible to know how much water is flowing through the filter, and thus know if it is according to the recommended values.

Despite the system supporting the existence of this sensor, at the time of writing, the MCU is only sending synthetic water flow data. The only thing left is the implementation code that allows the MCU to interact with the physical sensor.

4.1.7 Summary

By addressing these aspects in future work, the system can be further optimized for production-level deployment. The hardware adaptation, dynamic configuration, temperature compensation for pH measurements, and implementation of a hermetic enclosure will contribute to the system's robustness, accuracy, and long-term reliability in real-world applications.

Bibliography

- [1] D. Arias, "Adding salt to hashing: A better way to store passwords," Auth0. [Online]. Available: <https://auth0.com/blog/adding-salt-to-hashing-a-better-way-to-store-passwords/>
- [2] "Vorschriften," Mommertz Corporation. [Online]. Available: <https://mommertz.de/vorschriften.html>
- [3] "Mommertz," Mommertz Corporation. [Online]. Available: <https://mommertz.de/>
- [4] "ph," Wikipedia. [Online]. Available: <https://en.wikipedia.org/wiki/PH>
- [5] "Internet of things," Amazon. [Online]. Available: <https://aws.amazon.com/what-is/iot/>
- [6] "Microcontroller," Wikipedia. [Online]. Available: <https://en.wikipedia.org/wiki/Microcontroller>
- [7] "Web api," Wikipedia. [Online]. Available: https://en.wikipedia.org/wiki/Web_API
- [8] H. Cheung, "Measure ph with a low-cost arduino ph sensor board." [Online]. Available: <https://www.e-tinkers.com/2019/11/measure-ph-with-a-low-cost-arduino-ph-sensor-board/>
- [9] "Dht11 datasheet." [Online]. Available: <https://www.mouser.com/datasheet/2/758/DHT11-Technical-Data-Sheet-Translated-Version-1143054.pdf>
- [10] "Dht11-temperature and humidity sensor," components101. [Online]. Available: <https://components101.com/sensors/dht11-temperature-sensor>
- [11] B. Lutkevich, "Embedded system." [Online]. Available: <https://www.techtarget.com/iotagenda/definition/embedded-system>
- [12] "Sleep modes," Espressif. [Online]. Available: https://docs.espressif.com/projects/esp-idf/en/latest/esp32/api-reference/system/sleep_modes.html
- [13] "Esp32-s3 series datasheet," Espressif. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp32-s3_datasheet_en.pdf
- [14] "Esp32-s series datasheet," Espressif. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp32-s2_datasheet_en.pdf

- [15] "Esp8266ex series datasheet," Espressif. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/0a-esp8266ex_datasheet_en.pdf
- [16] "Esp32-c3 series datasheet," Espressif. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp32-c3_datasheet_en.pdf
- [17] "Msp430fr413x series datasheet," Texas Instruments. [Online]. Available: <https://www.ti.com/lit/ug/slau595b/slau595b.pdf?ts=1684329937122>
- [18] "Introduction to spring framework," Spring Framework Reference Documentation. [Online]. Available: <https://docs.spring.io/spring-framework/docs/3.2.x/spring-framework-reference/html/overview.html>
- [19] "What are the benefits of kotlin," Sagara Technology Idea Lab. [Online]. Available: <https://sagaratechnology.medium.com/what-are-the-benefits-of-kotlin-d7fdcd1cfc0>
- [20] "Time series database," Hazelcast. [Online]. Available: [title={https://hazelcast.com/glossary/time-series-database/},](https://hazelcast.com/glossary/time-series-database/)
- [21] "Relational databases vs time series databases," Influxdata. [Online]. Available: <https://www.influxdata.com/blog/relational-databases-vs-time-series-databases/>
- [22] J. Tao, "Time-series database basics." [Online]. Available: <https://devops.com/time-series-database-basics/>
- [23] "Db-engines ranking of time series dbms," DB-Engines. [Online]. Available: <https://db-engines.com/en/ranking/time+series+dbms>
- [24] "What is mqtt?" Amazon. [Online]. Available: <https://aws.amazon.com/what-is/mqtt/>
- [25] L. Dallinger, "Http vs mqtt: Choose the best protocol for your iot project." [Online]. Available: <https://cedalo.com/blog/http-vs-mqtt-for-iot/>
- [26] "Webpack documentation," Webpack. [Online]. Available: <https://webpack.js.org/concepts/>
- [27] "Field data management know to grow," CropX. [Online]. Available: <https://cropx.com/cropx-system/field-data/>
- [28] "Analog to digital converter," Expressif. [Online]. Available: <https://docs.espressif.com/projects/esp-idf/en/v4.4.1/esp32s2/api-reference/peripherals/adc.html>
- [29] "Non-volatile storage library," Expressif. [Online]. Available: https://docs.espressif.com/projects/esp-idf/en/latest/esp32s2/api-reference/storage/nvs_flash.html

- [30] "401 unauthorized," Mozilla. [Online]. Available: <https://developer.mozilla.org/en-US/docs/Web/HTTP/Status/401>
- [31] "Spring framework documentation - interface handlerinterceptor," springdoc-openapi v1.7.0. [Online]. Available: <https://springdoc.org/>
- [32] "Postgres - about," Postgres. [Online]. Available: <https://www.postgresql.org/about/>
- [33] J. Renders, "Why implementing a strong password policy is important for your business's data," brightlineIT. [Online]. Available: <https://brightlineit.com/why-implementing-a-strong-password-policy-is-important-for-your-businesss-data/>