Nov 14th, 1:50 PM - 2:15 PM

Secure Cloud-based IoT Water Quality Gathering for Analysis and Visualization

Soin Abdoul Kassif Baba M Traore
Kennesaw State University

Maria Valero
mvalero2@kennesaw.edu

Amy Gruss
Kennesaw State University

Traore, Soin Abdoul Kassif Baba M; Valero, Maria; and Gruss, Amy, "Secure Cloud-based IoT Water Quality Gathering for Analysis and Visualization" (2022). *KSU Proceedings on Cybersecurity Education, Research and Practice*. 12.
https://digitalcommons.kennesaw.edu/ccerp/2022/Research/12

Abstract

Water quality refers to measurable water characteristics, including chemical, biological, physical, and radiological characteristics usually relative to human needs. Dumping waste and untreated sewage is the reason for water pollution and several diseases to the living hood. The quality of water can also have a significant impact on animals and plant ecosystems. Therefore, keeping track of water quality is a substantial national interest. Much research has been done for measuring water quality using sensors to prevent water pollution. In summary, those systems are built based on online and reagent-free water monitoring SCADA systems in wired networks. However, centralized servers, transmission protocols, and data access can present challenges and disadvantages for those systems. This paper proposes a secure Cloud-based IoT water quality gathering architecture for water quality analysis and visualization to address the limitations of the current systems. The proposed architecture will send, analyze and visualize water quality data in the Cloud by utilizing specialized sensors and IoT-based gateways to capture water measurements (Dioxygen concentration, and temperature, among others). Then, they communicate securely to the Cloud-based server through a high-speed wireless network. We evaluated the performance of the proposed framework on a process-oriented approach to success metrics for cyberinfrastructures. The experiments were conducted in a laboratory and focused on network security and resiliency, the IoT prototype performance in dropping real-time data transmission, and remote access. The results demonstrate higher data collection and transmission effectiveness with minimal data loss and low energy usage over time. The accompanying cloud-based platform provided the flexibility needed for water quality monitoring and laboratory studies.

Disciplines

Computer and Systems Architecture | Environmental Engineering | Information Security | Management Information Systems | Technology and Innovation

Secure Cloud-based IoT Water Quality Gathering for Analysis and Visualization

1st Soin Abdoul Kassif Traore
Department of Computer Science
Kennesaw State University
Marietta, USA
straore3@students.kennesaw.edu
ORCID: 0000-0001-8913-9604

2nd Maria Valero
Department of Information Technology
Kennesaw State University
Marietta, USA
mvalero2@kennesaw.edu
ORCID: 0000-0001-8913-9604

3th Amy Gruss
Department of Environmental Engineering
Kennesaw State University
Marietta, USA
agruss@kennesaw.edu
ORCID: 0000-0002-2510-5808

***Abstract*—Water quality refers to measurable water characteristics, including chemical, biological, physical, and radiological characteristics usually relative to human needs. Dumping waste and untreated sewage are the reasons for water pollution and several diseases to the living hood. The quality of water can also have a significant impact on animals and plant ecosystems. Therefore, keeping track of water quality is a substantial national interest. Much research has been done for measuring water quality using sensors to prevent water pollution. In summary, those systems are built based on online and reagent-free water monitoring SCADA systems in wired networks. However, centralized servers, transmission protocols, and data access can present challenges and disadvantages for those systems. This paper proposes a secure Cloud-based IoT water quality gathering architecture for water quality analysis and visualization to address the limitations of the current systems. The proposed architecture will send, analyze and visualize water quality data in the Cloud by utilizing specialized sensors and IoT-based gateways to capture water measurements (Dioxygen concentration, and temperature, among others). Then, they communicate securely to the Cloud-based server through a high-speed wireless network. We evaluated the performance of the proposed framework on a process-oriented approach to success metrics for cyberinfrastructures. The experiments were conducted in a laboratory and focused on network security and resiliency, the IoT prototype performance in dropping real-time data transmission, and remote access. The results demonstrate higher data collection and transmission effectiveness with minimal data loss and low energy usage over time. The accompanying cloud-based platform provided the flexibility needed for water quality monitoring and laboratory studies.**

***Index Terms*—Water quality, cloud environment, IoT, security, minimal data loss.**

I. INTRODUCTION

Water is an essential requirement for life on our planet, so when the water in our rivers, lakes, and oceans becomes polluted, it can endanger wildlife, make our drinking water unsafe, and threaten the waters we swim in and fish [1]. Thus it is vital to keep track of the pollution level in the water body. Water quality is characterized by measuring several factors, such as the concentration of dissolved oxygen, bacteria levels, salt (or salinity), or the amount of material suspended in the water (turbidity).

Environmental Monitoring Systems (EMS) primarily gather data from various locations and provide the information to scientists for studies and policy-makers and planners for decision-making to solve ecological problems. Depending on the type of data needed, the materials used for EMS architecture can change, but the process remains the same for data transmission and exploitation. Water Monitoring systems are an essential part of EMS. They are based on SCADA systems in wired networks regarding water treatment stations. They often rely on satellite-based communication technologies for data transmission, which can be essential for remote locations but slow and costly otherwise [2].

Most of the existing SCADA systems store and process data in centralized servers, which display sensor values only in site locations, then the analysis process is sometimes more complicated. The systems lack sensors' scalability since most have programmable logic control (PLC) with analog/digital modules. Thus, it is hard to add new sensors. The data gathered by these sensors is very often transmitted without proper security methods to the SCADA server, which challenges data integrity and safety. Current systems have suffered multiple attacks on their data, and attackers do not hesitate to change the level of chemicals in water plants, such as in Florida in recent years [3].

The Internet of Things (IoT) and Cloud-based IoT offer a valuable platform for gathering, share, and analyzing data from remote places. Cloud-based IoT is a cost-effective and reliable solution for many sensor applications [2]. However, due to the sensitive data transmitted to the Cloud, a robust, secure platform must be implemented to avoid attacks on critical national infrastructure and data[4]. To overcome the above difficulties, this paper presents a secure Cloud-based IoT water quality gathering architecture that enables environmental sensors to gather local information from a water source and visualize it in real-time. The data of the integrated IoT devices is then preprocessed and secured before transmission using a wireless network e.g. (4G, 5G, Wi-Fi). Finally, the data is sent to a Cloud system composed of time-series databases and a visualization mechanism for posterior analysis [5]. It also provides remote real-time sensor access for calibration and error fixing.

This project has been founded with SunTrust Fellowship Program through the Institute of Cybersecurity Workforce Development.

The rest of the paper is organized as follows: Section II discusses Sensor Cloud Based Architecture related work of wireless sensor networks (WSN). Section III explains the approaches selected to solve the challenges. Section IV discusses the proposed Define architecture for Cloud-based IoT Water Quality Gathering; the materials, methods, and implementations. Finally, the experimental setting and results are described in Section V with conclusions.

II. RELATED WORK

Wireless sensor network (WSN) has become the most crucial field in networking due to their numerous applications and integration with the Internet of Things [6]. WSN-based applications for environmental monitoring have been implemented for applications such as water quality monitoring, water chemical monitoring, hydrodynamic performance monitoring, irrigation, and agriculture [7]. Furthermore, based on the application and availability of the energy resources, they can be deployed in areas with challenges such as temperature or humidity. WSN is a collection of sensor nodes organized into a cooperative network. Sensors in nodes have a short distance communication range; they run on low energy due to their small size, and they have to blend in their environment with minimal perturbations. Sensor nodes design consists of four central units that work as a whole, as presented in Figure 1; those components:

- The sensing unit is composed of environmental sensors that measure characteristics such as temperature, humidity, concentrations, pressure, etc.
- The processing unit, which is the brain of the node, does all processing and uses encryption schemes and protocols necessary for transfer. It is usually an electronic board or a microcomputer.
- The communication unit used a radio range to communicate with other nodes in the system.
- A power unit is usually a battery that provides other units with energy to run correctly.

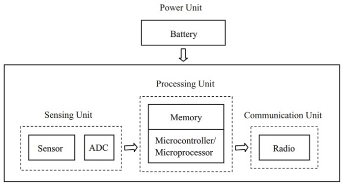

Fig. 1. WSN Basic Components

Powering WSNs can be challenging because of the finite amount of power a battery can hold. A power outage can be critical in an aggressive environment. Thus, it is important during the development process to use communication and routing protocols and algorithms that have a low impact on energy and do not affect sensor performances [8]. General purpose protocols such as Bluetooth and ZigBee are ideal for short-distance communication for a sensor to a sensor or a sensor to a computation unit[9]. Zakaria et al. work on a cloud-based WSN to monitor pH, conductivity, and dissolved oxygen parameters from wastewater discharged into water sources [10]. The authors designed a complete platform with a combination of General Packet Radio System (GPRS) for internet connectivity and a Global System for Mobiles (GSM) for a Short Message Service (SMS) system that provided a real-time IoT monitoring system to identify different water pollution events. IoT and remote sensing technology was applied in the work of Prasad et al.[11] to implement a Smart water quality monitoring system that measured Potential Hydrogen (pH), Oxidation and Reduction Potential (ORP), Conductivity, and Temperature parameters for water quality. Their design used an ADC, a GSM module, and a microcontroller; the transmission protocol was FTP.

III. APPROACH

This paper is an interdisciplinary collaboration between the Department of Civil and Environmental Engineering and the Department of Computer Science at Kennesaw State University (KSU), which will primarily use the research framework. This department provided feedback about the best IoT sensors that produce significant results in their research on wastewater plants. From this inquiry, we decided to collect three main water parameters:

- Temperature is applicable in behavioral analysis of the parameters being measured.
- TDissolved Oxygen concentration that is important in research for wastewater plants in the biological process [12].
- Saturation.

After determining the suitable types of sensors, we compared different products on the market with the desired measurements and their compatibility with our architecture. We selected the *Go Direct Optical Dissolved Oxygen Probe* after extensive research [13].

One key factor to consider in our solution was the environmental challenges that limited the type of computation unit we could choose. The computing unit will receive data from the sensor through Bluetooth, process it, and send it to cloud-based servers. Therefore, we must choose an energy-efficient unit with WiFi capability, SD card storage with monitoring systems, and remote access built-in. With all other abilities, Raspberry Pi (RP) is well suited to have remote access through Secure Shell (SSH) or Virtual Network Computing (VNC). In addition, it is bundled with a Liquid Crystal Display (LCD) screen to track different processes the device has to go through before retrieving sensor data. Besides those, multiple Raspberry Pis energy consumption studies will be helpful during evaluation [14]. Also, RPs provided enough computational power for encryption schemes and protocols.

Then, we consider the data transmission and accessibility surely through the Cloud. A Cloud-based time series database to store the data is hosted on campus and only accessible through a Vitual Private Network (VPN), which does not work well on RPs. We opted for a secured SSL/TLS (Secure Socker

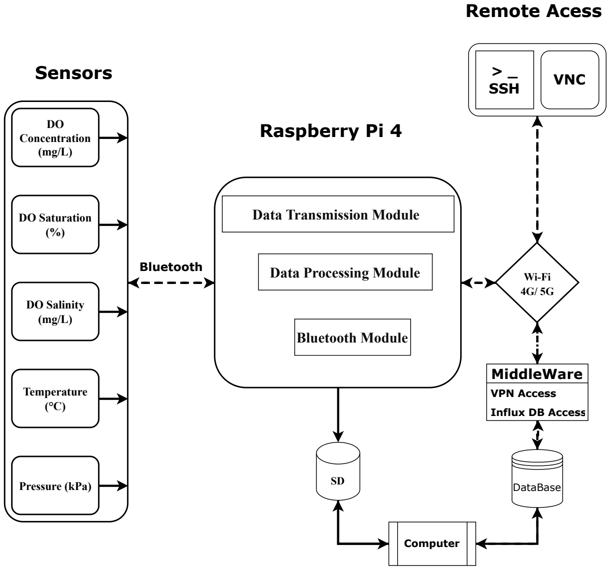

Fig. 2. Proposed Framework Architecture

Layer / Transport Layer Security), coupled with HTTPS (Hypertext Transfer Protocoal Secure) for data transfer and access through the platform [15]. Encryption required much computational power, so we chose security protocols with minimal CPU usage. The communications will pass through a middleware securely accessible via APIs for an additional security layer.

Finally, a web-based data visualization system using Grafana [16] is provided for the sensor and IoT device monitoring purposes, where we could monitor the RP CPU, temperature, memory, and other critical data to ensure good performance of used devices. Also, the same platform is used by other departments to access data collected on-site for their studies or analysis purposes. All data are retrieved directly from the cloud-based server located on the KSU campus.

IV. PROPOSED ARCHITECTURE AND IMPLEMENTATION

The proposed architecture is based on a WSN architecture described in Figure 2, designed with five sensors and a computational unit with the communication layer embedded. Thus, we have a physical section composed of a RP4 and a Probe embedding sensors that communicate with a cloud-based platform with the device monitoring part and several configured servers for data treatment. Those two sections use different wireless networks available in the wastewater plant to communicate. The end-point user can access the data through a web link on a PC or a mobile.

A. *Sensors and Raspberry Pi*

The system has to measure the dissolved Oxygen (DO), defined in biological treatment as the relative measure of oxygen dissolved in wastewater available to sustain life, including living bacteria [12]. In addition, the system has to measure the temperature simultaneously, as explained in the sections above. We used a probe with five sensors for DO concentration, water temperature, and atmospheric pressure. The probe gathers data and transfers it to the RP4 through Bluetooth without human intervention, using packages provided by the manufacturer that drives the probe. The process is handled by an algorithm we developed to implement the platform and connect all different sections. RP4 also comes with proprietary software with limited access and approximately 24 hours of battery life on a single charge [17].

B. *Software configuration*

Multiple software intervene on different levels of the system architecture. We developed a program that gives us the scalability and flexibility necessary to control our IoT platform. The program has three primary purposes implemented using different Python libraries and modules, as shown in Figure 3.

The first level of the program aims to connect to the probe and give us feedback on the process displayed by the LCD screen. This part starts the LDC screen, which displays greeting messages, and the date and time the program start. Next, the Bluetooth module is initialized for instruction from the Vernier packages, which will send product information for connecting to the probe. The program moves to the next part when this process is successful.

The second level retrieves the data from the sensors embedded in the probe. When a collection starts, the probe transmits the data in a maximum 1 X 5 array corresponding to the number of sensors detected. A function parses and labels the retrieved data before saving a copy in a text file on the RP4 SD card and passing another copy to the next level. If an error occurs at this stage, a message is displayed on the LDC screen, and the program starts the closing process.

The third level is a loop which starts by sending a ping to the cloud-based server and waiting for a response. Then depending on which network the device has access to, an API tries to open a secure communication channel with middleware or uses InfluxDB [18] API to directly with the Cloud database if there is a VPN installed on the Raspberry PI. In case of an error during this step, the program exits the loop and returns to the previous level. After a successful handshake process and SSL/TLS protocols to establish a secured channel, the program starts the data encapsulation process using a JSON module and sends it through the secured channel. At this level, the program relays heavily on a Cyberinfrastructure created using the Cloud for this paper.

C. *Cyberinfrastructure*

Computation power and energy efficiency are critical in IoT research, so implementing a system that relays on multiple security and encryption schemes can take our designs to their

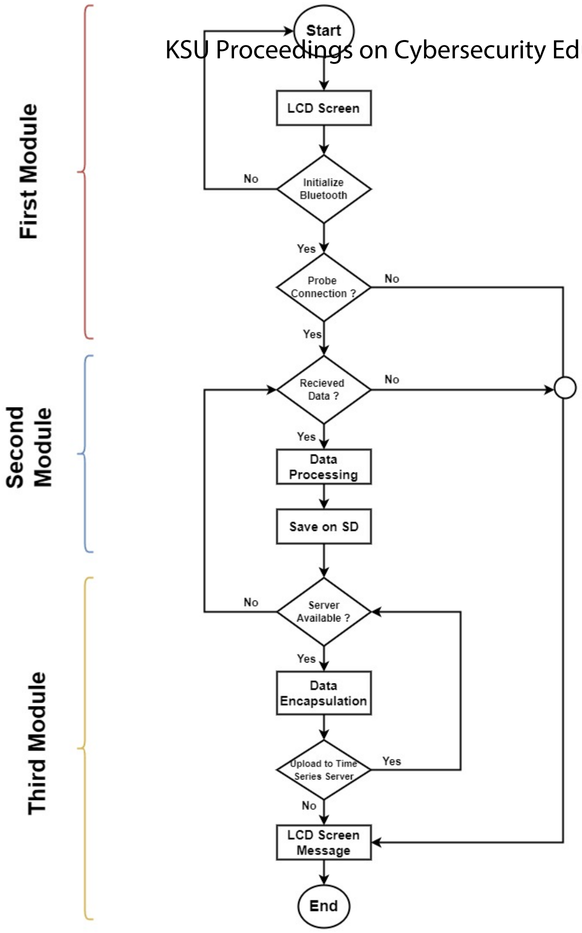

Fig. 3. Flowchart of the framework Implemented Program

limits. Thus we decided to add an API Middleware layer that provides a bridge between client and backend systems [19] to be able to communicate with the university server at any time. The framework uses InfluxDB, a time series database designed to store extensive data with a quick real-time performance well suited to IoT sensor data [20]. We also used Grafana, an open-source web application that offers multi-platform analytics and interactive data visualization and is scalable with real-time updates via HTTP API and InfluxDB connections API.

V. EXPERIMENTAL RESULTS AND DISCUSSIONS

We started by testing the IoT device in the lab with no access to the internet. Everything runs on a Linux desktop with no secured protocol or encryption using the desktop IP address hardcoded in the program. We used tap water to test the framework for 2 hours. We set a local Docker engine with Grafana hosted on port 3000 and InfluxDB hosted on port 8086 containers. The real-time reading of the sensor showed only temperature variation during the experiment.

The second measurement step tested the framework access to the campus resources using a VPN. Again, the goal was to securely transfer data to the InfluxDB server and monitor it in real-time on a Grafana dashboard. For this experiment, we used tap water. Figure 4 show the sensors, and Figure 5 shows the Grafana dashboard second experiment over one-hour results.

Fig. 4. First Lab Experiment

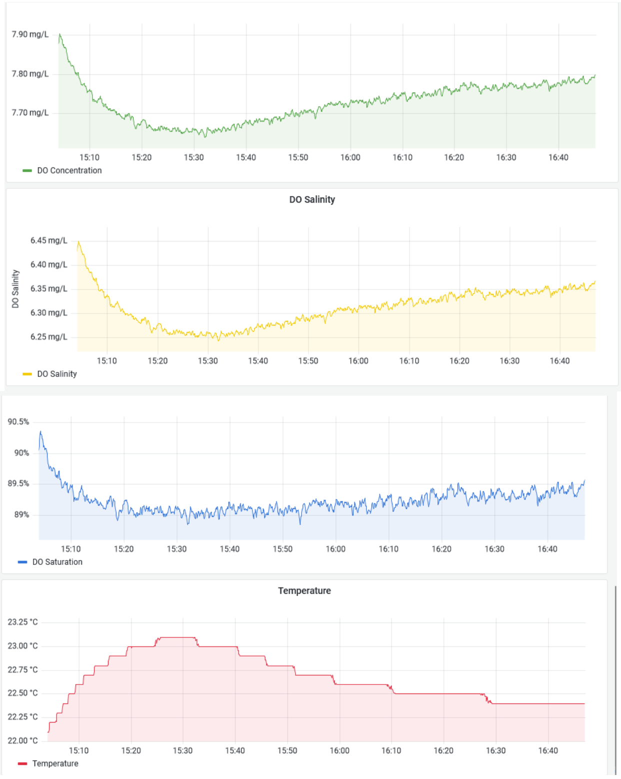

Fig. 5. Grafana dashboard Second Experiment

The third experimental set was in an Environmental Lab with wastewater in a proper setting. It was the most extended experiment conducted in two days, from September 1st to September 3rd, 2022, without human intervention. We did not use any battery for this experiment, and the probe was fully charged before the investigation. We were interested in the variation of DO during the experiment. Figure 6 and 7 shows a significant variation of DO over time when the experiment is conducted with wastewater. The IoT device was presented to the environmental engineering course Lab, and students used it during experiment installation. The monitoring dashboard shows an offline period of four hours due to a Wi-Fi outage. The missing data were recovered from the local text file storage. Following this experiment, we updated the platform program to handle better Wi-Fi shortages caused by the campus network security protocols. Considering the promising results of these first phases of experiments, we can affirm that the platform implements our proposed architecture. We plan to start the second phase of testing, which will be done using batteries in the lab and wastewater treatment plant. We will get after those experiments more data on water quality and more data on our platform for an extensive evaluation.

Fig. 6. Grafana dashboard Third Experiment Experiment

Fig. 7. Grafana dashboard Third Experiment

VI. CONCLUSION

The efficiency of IoT-based platforms makes them the perfect tools for low-cost and high-performance research studies. In contrast to over-ready tools, IoT frameworks are robust because they use open-source material with well-supported communities behind them. The rapid development of IoT platforms can be a challenge when it comes to securing them. Water quality measures can be challenging because of moisture and remote location access. Thus it is essential to have a reliable system with access to the internet and remote access to the IoT device used so minor error, and bugs can be resolved online. On the other hand, it is critical to select proper data transmission protocols that minimize the energy usage of the platform and provide good data quality and integrity. This paper presents a Secure Cloud-based IoT Water Quality Gathering framework implementation from scratch and provides the framework preliminary test. We are currently working on the following experiments to prove the efficiency of such performance.

REFERENCES

[1] U. Hölzle and et al., "The datacenter as a computer – an introduction to the design of warehouse-scale machines."
[2] P. Snyman, "Risk-based capital measures for operational risk management," 2011.
[3] 2022. [Online]. Available: https://www.securitymagazine.com/articles/94552-hacker-breaks-into-florida-water-treatment-facility-changes-chemical-levels
[4] S. Ahmed and S. Ismail, "Water pollution and its sources, effects & management: A case study of delhi," 2018.
[5] Z. Rasin and M. R. Abdullah, "Water quality monitoring system using zigbee based wireless sensor network," 2020. [Online]. Available: http://eprints.utem.edu.my/id/eprint/8648/1/91410-7575_ijet-ijens.pdf
[6] G. Verma and V. Sharma, "A novel thermoelectric energy harvester for wireless sensor network application," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 5, pp. 3530–3538, 2019.
[7] A. T. Demetillo, M. V. Japitana, and E. B. Taboada, "Sustainable environment research," 2019. [Online]. Available: https://sustainenvironres.biomedcentral.com/articles/10.1186/s42834-019-0009-4
[8] G. Yin, G. Yang, W. Yang, B. Zhang, and W. Jin, "An energy-efficient routing algorithm for wireless sensor networks," in *2008 International Conference on Internet Computing in Science and Engineering*, 2008, pp. 181–186.
[9] D. A. Bashar, "Sensor cloud based architecture with efficient data computation and security implantation for internet of things application," 2020. [Online]. Available: http://eprints.utem.edu.my/id/eprint/8648/1/91410-7575_ijet-ijens.pdf
[10] Y. Zakaria and K. Michael, "An integrated cloud-based wireless sensor network for monitoring industrial wastewater discharged into water sources," *Wireless Sensor Network*, vol. 09, pp. 290–301, 01 2017.
[11] A. N. Prasad, K. A. Mamun, F. R. Islam, and H. Haqva, "Smart water quality monitoring system," *2015 2nd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, pp. 1–6, 2015.
[12] D. Theobald, "Stackpath," 2013. [Online]. Available: https://www.watertechonline.com/wastewater/article/15543520/what-are-process-controls-for-dissolved-oxygen-during-biological-treatment

[13] Y. Wei, Y. Jiao, D. An, and D. Li, "Review of dissolved oxygen detection technology: From laboratory analysis to online intelligent detection," 2019. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6767127/

[14] G. Bekaroo and A. Santokhee, "Power consumption of the raspberry pi: A comparative analysis," 08 2016, pp. 361–366.

[15] F. Van den Abeele, T. Vandewinckele, J. Hoebeke, I. Moerman, and P. Demeester, "Secure communication in ip-based wireless sensor networks via a trusted gateway," *2015 IEEE 10th International Conference on Intelligent Sensors, Sensor Networks and Information Processing, ISSNIP 2015*, 05 2015.

[16] Grafana Labs, "Grafana," 2018. [Online]. Available: https://grafana.com/

[17] 2021. [Online]. Available: https://www.vernier.com/manuals/gdx-odo/

[18] Influxdata Inc., "InfluxDB," 2019. [Online]. Available: https://www.influxdata.com/

[19] O. Harfoushi and R. Obiedat, "An insight to middleware technologies used for cloud computing frameworks: From the perspective of successful enterprise application integration," *Modern Applied Science*, vol. 12, p. 234, 08 2018.

[20] S. Gangadhar S, "The real time environmental time series data analysis using influx db," vol. 1, Dec. 2020. [Online]. Available: https://www.ijasi.org/index.php/ijasi/article/view/4