Journal of Theoretical and Applied Information Technology
31st December 2021. Vol.99. No 24

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A PROPOSAL ARCHITECTURE FOR DATA GATHERING AND PROCESSING IN INDUSTRY 4.0

[1]OUAHIB AJIL, [2]MAHA REZZAI, [3]WAFAA DACHRY, [4]HASSAN GZIRI, [5]AHMED MOUSRIJ

[1,3,4,5] Laboratory of Mechanical Engineering, Industrial Management and Innovation, Hassan First University of Settat, Morocco

[2] C3S, Higher School of Technology, Hassan II University of Casablanca, Morocco

E-mail : [1] ouahib.ajil@gmail.com, [2]maha.rezzai@gmail.com, [3]wafaa.dachry@gmail.com, [4]hassan.gziri@gmail.com, [5]mousrij@gmail.com

ABSTRACT

The deployment of the Industrial Internet of Things (I-IoT) in Industry 4.0 aims to enhance the productivity by gathering and processing of real-time data. Industry 4.0 uses Big Data ecosystem to deal with the large volume of heterogeneous data generated by several types of sensors. Therefore, Big Data technologies offer scalable IT resources for computing and storing. However, there is a scarcity of studies on data flow platforms to address the core challenges of the I-IoT complex environment. This paper proposes a generic platform for data gathering and processing from the sensors to the cloud, and using Edge computing approach that introduces an intermediate layer between the Cloud and the I-IoT nodes. Vitals features are deployed in this layer such as security management, Stream Processing, local storing in order to ensure real time monitoring and to reduce the latency in case of a poor Internet connectivity.

Keywords: *Industry 4.0, I-IoT Data Gathering and Processing, Big Data, Edge Computing, Stream Processing, Cloud.*

1. INTRODUCTION

The fourth industrial revolution, named Industry 4.0, and the advent of the new concept of the Industrial Internet of Things (I-IoT) have exponentially increased the need for data management from gathering to processing.

Industry 4.0 relies on the use of data for better production management through the development of intelligent modules able to understand, predict and act in real time [1]. Note that industrial data has specific characteristics such as large volume, variety, variability and velocity. For this reason, the fourth industrial revolution is closely linked to the Big Data ecosystem. The latter brings several new technologies described as follows [2]:

- Hadoop Distributed File Systems (HDFS) to handle large amount of data.
- Data Lake to store all structured and unstructured data at any scale.
- Big Data Analytics and machine learning to make insights and extract valuable information.

On the other hand, I-IoT is the extension and use of the Internet of Things (IoT) in the industrial domain[3].

Unlike the IoT, the I-IoT environment is more complex and consists of heterogeneous technologies including different communication protocols [4], several types of sensors that generate variable data at multiple frequencies and format [5]. In addition, these sensors are usually physically deployed in conditions that considerably influence the reliability of the data they generate.

Furthermore, there is a lack of studies and literature review to document the data flow architecture design and implementation [6]. In addition, most of the commercial platforms are not open source and include proprietary technologies, which limit portability and interoperability [7]. Thus, all those factors accentuate the necessity of developing a generic and open source platform to address several challenges such as security, data interchange formats and interoperability.

The present paper proposes a generic architecture for structured and unstructured data gathering and processing. This architecture is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

designed independently of communication technologies, able to connect and communicate in real time with various heterogeneous devices of the I-IoT. It also provides cybersecurity, data reliability and filtering, diagnosis, prognosis and predictive maintenance based on open source BIG DATA and Machine Learning technologies. All of those features are deployed on three separated layers, "Factory" layer for data gathering, "Edge" layer for stream processing and "Cloud" layer for Big Data ecosystem.

2. RELATED WORK

In this section, we exhibit and discuss some representative platforms in the area of data gathering and processing and fog computing in manufacturing.

Emeakaroha et al. [7] propose and implement a generic architecture for a Cloud based sensor and IoT-device data monitoring, gathering and processing platforms. Their proposed platform aims to provide an open solution implementing standardized technologies to promote fast adoption. However, the authors did not take into consideration the implementation of fog computing, which includes vital modules that ensure production continuity and data backup in case of Internet outage.

Moreno et al. [8] propose a holistic IoT-based management platform for smart environments. The authors take into consideration the rapid growth and deployment of sensors and smart devices and the heterogeneity of the data they generate. The proposed architecture aims to address the interoperability issues in gathering, processing IoT data and actuating control.

Lee at al. [9] present a 5C architecture for Cyber-Physical Systems in Industry 4.0 manufacturing systems. The paper provides a viable guideline for developing and deploying a CPS for manufacturing application. The work focused on defining, through a sequential workflow manner, how to construct a CPS from the initial data acquisition, to analytics, to the final value creation.

Yu et al. [10] present an integrated CPS based architecture for smart manufacturing and provide the deployment details, addressing all the potential problems in an appropriate way. The authors described the working principle through three parts, namely big data ingestion, big data management, and big data analytics. Their architecture could be easily used and extended in a smart manufacturing like predictive maintenance system.

Luis de Moura et al. [11] propose an agnostic federated management architecture that meets any IIoT devices independent of the technology specificities or communication mechanism. The proposed architecture builds a flexible structure to support the management of any device, allowing a holistic view of all devices available in industrial operations. However, the proposed architecture does not take into account the deployment of machine learning algorithms for intelligent management as well the transfer of data to the cloud.

Zhang et al. [12] present a feasible method for designing an autonomous factory with exception-handling capabilities. The goal is to quickly organize the production, discover and deal with abnormities without human intervention. Three main contributions presented in this work. Firstly, the architecture and the function models of the intelligent shopfloor provides a reference for the future designs. Secondly, a cyber-physical system for manufacturing shopfloor based on the multiagent technology is developed. Thirdly, the self-organizing and self-adaptive mechanisms are introduced. However, The insufficiency of the case analysis presents one of the main limitations of this study.

Boyes et al. [3] propose an approach for analysing the city level risks and vulnerabilities to inform both system planning and design. The authors developed a comprehensive analysis methodology that offers a systematic way to study the cyber–physical systems and identify safety, security or resilience issues that need to be addressed in the systems design or operation.

Wu et al. [13] introduce a novel architecture that enables remote real time sensing, monitoring, and scalable high performance computing for diagnosis and prognosis. Their architecture utilizes wireless sensor networks, cloud computing, and machine learning. The objective is to demonstrate how their solution can enable manufacturers to monitor machine health conditions and generate predictive analytics by using a fog computing-based framework for data-driven machine health and process monitoring in cyber-manufacturing.

Wang et al. [14] propose a new platform named fogIBDIS that integrate and share industrial Big Data with high raw data security and low network traffic loads. The proposed platform applies the data processing in the fog clients within the manufacturing systems, thereby changing the centralized data-processing mode into distributed task execution.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. CHALLENGES AND OBJECTIVES

3.1 Challenge

The I-IoT environment is more complex and specific, which presents challenges in designing a robust architecture to gather data. The first challenge is to deal with the heterogeneity of the acquisition devices by finding the right way to let equipment communicate and work together, regardless of its manufacturer or role. The second challenge is related to the variety of the gathered data, its volume and its velocity. While the third challenge concerns the quantity of noise that may be collected during the acquisition of data. Thus, data must be cleaned and filtered before the storing process. Other challenges may be added to the list like dealing with Internet outages and slowness, real time monitoring, security and data access control.

3.2 Objectives

Based on the articles presented in the literature regarding data flow in Industry 4.0, we have crossed and analyzed the different criteria aiming to deploy the appropriate technologies and the adequate schema of data flow management in order to address the different challenges mentioned above. Besides, we aim to deliver the following list of criteria:

- Open-source solution to provide flexibility, extensibility and customization.
- Interoperability manager to ensure horizontal and vertical communications and to deal with the heterogeneity challenge.
- BIG DATA and Cloud technologies for storing and processing to guarantee scaleability of the system and providing the necessary computing and storage resources. As a result, our architecture will handle the quantity, variety and velocity of the gathered data.
- Edge Computing that host vital features to ensure real time data analysis and local storage to avoid data loss.
- Security management.
- Machine Learning algorithms to realize efficient and deep analysis for diagnostic, prognostic and predictive maintenance.
- Rules-based expert system to rise the speed of analysis.

The table 1 resumes the evaluation of the main criteria and compares our proposed work to other similar architectures from recent literature.

To address those objectives and challenges, we adopt a data flow architecture as presented below.

3.3 High-level Data Flow

Figure 1 illustrates the data flow of the proposed architecture and shows the principle of our design. Indeed, our solution consists of splitting the flow into three layers: "Factory layer", "Edge layer" and the "Cloud layer". Each layer has specific features:

3.3.1 Factory

The factory refers to the physical world, the world of sensors, actuators controllers and so on. This layer performs the following functionalities:

- Read the measurements that arrive from sensors.
- Convert analogic signals to numerical values.
- Ensure interoperability by deploying an OPC UA server.
- Controlling, supervising, planning and scheduling the entier production chaine.

3.3.2 Edge

The edge is the layer that enables the I-IoT. It is physically located in the factory and is linked from one side to the data sources and from the other side to the cloud.

As shown in the figure 1, the edge must have the following features:

- Implement the most common network and Internet communication protocols and technologies such as Message Queuing Telemetry Transport (MQTT), Hypertext Transfer Protocol Secure (HTTP), ZigBee, Advanced Message Queuing Protocol (AMQP), OPC UA Client interface to:
  - Gather data from the industrial data sources.
  - Transfer data to the cloud.
  - Receive commands and configurations from the cloud.
- Formatting, cleaning and sorting the gathered data.
- Stream data analysis for the survey, diagnosis and prognosis process.
- Implement a local database and apply a store and forward mechanism to avoid data loss in case of a poor or unstable connectivity.
- Guarantee the cybersecurity.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.3.3 Cloud

The power and the flexibility of the cloud ecosystem are essential for BIG DATA [15]. This ecosystem allows computing resources to be scaled up or down as demand changes.

The most relevant components managed in the cloud are the following:

- Batch data processing for Big Data Analysis and Predictive Maintenance.
- Data Lake, Digital Twin and visualization technologies.
- Machine Learning.

The Cloud is also essential for gathering external data such as data from other production sites, from suppliers and from customers.

In the next section, we describe step by step the data flow of our proposed platform architecture, and we specify the role of each manager.

4. PROPOSAL ARCHITECTURE

In this section, we present and detail our proposed architecture and its components.

The obvious first step is to extract data from the physical world and convert the analogic signals into digital data. This step takes place in the CIM Pyramid and is carried out by the sensors and controllers as shown in figure 2. This brings us back to define the first block in our architecture.

4.1 CIM Pyramid

Computer Integrated Manufacturing, or CIM, is a logical and hierarchical model for production systems. Developed in 90s, it is a reference model for the implementation of industrial automation based on the collection, coordination, sharing and transmission of data and information between the different systems and subsystems by means of software applications and communication network.

The CIM model is depicted as a pyramid made up of five functional levels as shown in figure 2:

- Field levels: This level includes all devices that interact directly with the process, such as sensors.
- Control System level: This level includes the devices that control directly the activity and the state of the sensors and actuators such as PLC or microcontrollers.
- Supervisory level: The main function of this level is to exchange data and instructions between the lower and the upper levels. Indeed, data collected from lower levels and transferred to the upper ones and instruction are receipted from upper levels, transformed to actions and commands to the lower levels.
- Management Level: This level concerns the order management and the production planning. It can also be named Manufacturing Execution System.
- Company Management: It includes the systems and software packages that use companies to manage all their activities and business, such as human resources, project management, accounting, etc.

After the data extraction step, the next module has a key role in unifying communication protocols:

4.2 OPC UA Server

Our proposed architecture relies heavily on the OPC-UA server in order to ensure and facilitate interoperability for horizontal communication between machines and also for vertical communication between machine and “Edge”.

Indeed, Open Platform Communications Unified Architecture (OPC-UA) is a specification-based standard developed in collaboration between manufacturers, research centres and users in order to facilitate consistent information exchange in heterogeneous systems [16].

Thus, this server extracts the data from the various sensors and controllers and routes it continuously to the second block of our architecture, which is the "Edge".

4.3 Edge Computing

4.3.1 Communication Manager

Communication manager is a central module responsible for exchanging data and instructions between the different managers. The main features of this manager are as follows:

- Includes an OPC UA Client to communicate with the OPC UA server and gather data from it.
- Collect useful information from the high levels of the CIM Pyramid that concern management and scheduling.
- Transfers the local database to the Data Lake on the cloud.
- Forwards Machine Learning API from the Cloud to the Edge for Stream Data Analysis.
- Routes the results of stream data analysis to cloud for maintenance and BIG DATA Analytics.

Once the data is collected, the next manager convert it into a useful format.

4.3.2 Data Formatting Manager

The first step of data pre-processing is to determine the appropriate format in which each type of data must be stored and used. Indeed, the Data Formatting Manager deals with different brands and types of sensors that produce different types of data. In addition, the formatting process must be done without a significant loss of information. For example, this Formatting Manager should convert a file generated by a sound sensor from wave extension to mp3 extension and maintain the same quality and reliability of the collected sound.

4.3.3 Stream Data Analysis

At this step, we introduce the four managers of the Stream Processing:

4.3.3.1 Data Filtering Manager

Filtering data is the step of the quality control of the gathered data. In fact, this manager completes the controls made in the CIM Pyramid by adding deeper analysis based on Machine Learning algorithms. Those analyses take into consideration all the parameters that can influence the accuracy of the sensors such as temperature, humidity, pressure of the production environment. During this step, the module filters out the captured errors and corrects them. That will help to make all further analysis and processes (diagnostic, maintenance, modelling, etc.) more efficient. At the same time, it is important to flag erroneous data and report them as such in order to determine the origins of these errors.

4.3.3.2 Survey Manager

The purpose of this manager is to monitor and check the status of all devices: Their health, efficiency and performance. It has a descriptive role and answers the question of what is happening. Indeed, it monitors and analyses live data to give a complete overview of the current status.

4.3.3.3 Diagnosis Manager

The diagnostic manager is responsible for identifying the origins of anomalies and failures in the working environment, including sensor inaccuracy or malfunctions, mechanical and electrical failures of production machines and other devices. It gives us an answer to why is it happening. This function is made possible by the log files created by the devices themselves and by Machine Learning algorithms using the data stored in the local database.

4.3.3.4 Prognosis Manager

The Prognosis Manager provides answers about what will happen. It has a preventive purpose and gives insights of future faults and anomalies with their risk of impact and the probability of occurrence. In the same way as the Diagnosis Manager, it uses Machine Learning algorithms such as Deep Learning, Regression and Data Driven. It is an essential asset for successful predictive maintenance.

4.3.4 Rule-Based Expert

Rule-based Expert gives an assistant to Stream Data Analysis in their functionalities. It contains prescribed knowledge to solve problems. In fact, this manager has a register that contains all information about the different machines, components, controllers, sensors and even software and operating systems. For example, this registry records a detailed list of the specificities of a sensor, such as its life cycle, calibration parameters, operating conditions, error codes; list of contacts, etc. the combination of this knowledge and the Machine Learning promise more effective analysis.

4.3.5 Configuration Manager

This manager is responsible for the updates, patches and configurations of the different hardware and software. However, each task must be planned by the "maintenance manager" to be able to execute them without stopping the production line.

4.3.6 Storing Manager

The filtered data must be stored in databases that make it conducive for machine learning. In addition, the types of databases (DB) chosen must allow instantaneous reading/writing. Thus, and given the huge amount of data generated by sensors and their varieties, the new generations of NoSQL databases dedicated to BIG DATA are the most recommended to satisfy those requirements. In this perspective, the Storing Manager routes the filtered data to temporary local databases and set up a store and forward mechanism to avoid data loss in case of a poor or unstable connectivity.

Databases are transferred to the cloud and stored on a DATA Lake which can be queried for further BIG DATA Analytics like visualization, modelling, etc. Note that the frequency of transfers and the duration of temporary databases are parameters to be adjusted according to requirements and the IT resources.

4.3.7 Security Manager

I-IoT devices present many vulnerabilities and privacy risks, both internally and externally. This manager defines the necessary rules and manages the various tools to block these attacks and thus reach and maintain cybersecurity according to standards. These features include:

- Installing a firewall to block direct exposure of devices to the Internet.
- Executing security patches on every system and program.
- Auditing all network traffics.
- Managing access controls, authorizations and authentication modes.

EDGE functionality ends at this point; the data is then transferred to the cloud to perform the processing that requires more storage and computing resources.

4.4 Could Computing

As for the "Edge computing", a communication manager gathers data from the edge and store it on a Data Lake.

- Data Lake is a centralized repository that allows us to store all structured and unstructured data at any scale [17].
- As the data is not sent continuously, our architecture adopts Batch Processing technologies to perform other vital features such as:
- BIG DATA Analytics: A set of tools dedicated to the exploration and analysis of data stored in the DATA Lake in order to extract relevant information and insights.
- Data Visualization: A set of methods for graphically summarizing data. It is part of data science and BIG DATA and considered as an essential step for the proper understanding of data.
- Digital Twin: Considered as one of the key technologies of Industry 4.0. The digital twin is a mirror of the physical world. In contrast to the Digital Shadow limited to the visualization, the Digital Twin gives us numerous advantages such as performing simulations, remote controlling, monitoring, etc.

The Cloud also includes other managers as below:

4.4.1 Maintenance Manager

The maintenance manager predicts and adjusts maintenance intervals to optimize planning and scheduling. It gives an answer the question what to do and when to do it, it has a perspective role.

It should be noted that predictive maintenance is one of the most widely discussed topics in Industry 4.0.

To this end, this manager exploits the results of the Diagnosis and Prognosis Manager, retrieves data from the Rule-based Expert and Configuration Manager, also gathers planning information from the CIM pyramid and gives insights about the action that needs to be taken and when to be taken.

Among the technologies used for this purpose are Deep Learning, Reinforcement Learning, Simulations on the Digital Twins, etc.

4.4.2 Machine Learning Manager

This manager includes algorithms used in data modelling. It is composed of three sub-modules:

- Model Builder: Building is the process of creating a machine learning or deep learning model using an appropriate algorithm with huge volume of data.
- Model Deplorer: makes such a Machine Learning or Deep Learning model available to Stream Data Analysis by converting the builder models into an Application Programming Interface (API).
- Model Monitoring and Assessment: ensures continuous performance quality checks on the deployed models.

5. DISCUSSION

The proposed architecture in this paper highlights several important concepts that concern I-IoT and Data Processing. Here are some advantages of our design over others:

- Rely only on open source technologies to allow plugging in new applications and to avoid the limitations of specific technologies
- Show the ability to connect and extract data from heterogeneous platforms.
- Use Edge Computing to ensure real time monitoring and control as well as to avoid data loss in case of Internet connection disruption by adding a local database.
- Improve the Stream processing by adopting a hybrid approach that combines Machine Learning algorithms and rule-base expert system.

- Take advantage of Cloud technology and Big Data for better management of IT resources in terms of computing and storage.
- Combine Stream Processing and Batch Processing to embrace the potential of both of them and to extract the maximum of useful information from the gathered data.

6. CONCLUSION

In this paper, we have presented our architecture platform for gathering and processing data. Its objectives such as using open-source technologies and deploying BIG DATA ecosystem, its components and its data flow. Our solution consists of splitting the data flow into three big layers including "Factory", "Edge" and "Cloud" layer. And also splitting processing into “Stream Processing” in the EDGE layer and to “Batch Processing” in the Cloud.

The strengths of the presented architecture comes from the utility of open source technologies in ensuring the customizability in various environment, the ability to gather data from heterogeneous devices and to combine different AI technologies such as machine learning and rule-base expert system to optimize the time of data analysis. However, the architecture does not takes into consideration the communication protocols that can help to reach optimally the real-time monitoring and control.

Note that this architecture will serve as a roadmap for the future work of our research laboratory and will be further developed and implemented in real conditions.

In our next steps, we intend to use Machine Learning algorithms to develop the Filter Manager of our proposed architecture. Then, we will test it with simulated data.

REFERENCES:

[1] S. Munirathinam, « Chapter Six - Industry 4.0: Industrial Internet of Things (IIOT) », in *Advances in Computers*, vol. 117, No.1 Elsevier, 2020, pp. 129-164. doi: 10.1016/bs.adcom.2019. 10.010.

[2] A. Oussous, F.-Z. Benjelloun, A. Ait Lahcen, and S. Belfkih, « Big Data technologies: A survey », *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 30, No. 4, oct. 2018, pp. 431-448, doi: 10.1016/j.jksuci.2017.06.001.

[3] H. Boyes, B. Hallaq, J. Cunningham, and T. Watson, « The industrial Internet of things (IIoT): An analysis framework », *Comput. Ind.*, vol. 101, oct. 2018, pp. 1-12, doi: 10.1016/j.compind.2018.04.015.

[4] M. Tabaa, F. Monteiro, H. Bensag, and A. Dandache, « Green Industrial Internet of Things from a smart industry perspectives », *Energy Rep.*, vol. 6, nov. 2020, pp. 430-446, doi: 10.1016/j.egyr.2020.09.022.

[5] E. A. Hinojosa-Palafox, O. M. Rodríguez-Elías, J. A. Hoyo-Montaño, J. H. Pacheco-Ramírez, and J. M. Nieto-Jalil, « An Analytics Environment Architecture for Industrial Cyber-Physical Systems Big Data Solutions », *Sensors*, vol. 21, No. 13, 2021, pp. 4282, doi: 10.3390/s21134282.

[6] X. Wu, V. Goepp, and A. Siadat, « Cyber Physical Production Systems: A Review of Design and Implementation Approaches », *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, Macao, 2019, pp. 1588-1592. doi: 10.1109/IEEM44572.2019.8978654.

[7] V. C. Emeakaroha, N. Cafferkey, P. Healy, and J. P. Morrison, « A Cloud-Based IoT Data Gathering and Processing Platform », *3rd International Conference on Future Internet of Things and Cloud*, Italy, 2015, pp. 50-57. doi: 10.1109/FiCloud.2015.53.

[8] M. V. Moreno, A. F. Skarmeta, L. Dufour, D. Genoud, and A. J. Jara, « Exploiting IoT-based sensed data in smart buildings to model its energy consumption », *IEEE Int. Conf. Commun. ICC*, 2015, pp. 698-703, doi: 10.1109/ICC.2015.7248403.

[9] J. Lee, B. Bagheri, et H.-A. Kao, « A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems », *Manuf. Lett.*, vol. 3, 2015, pp. 18-23, doi: 10.1016/j.mfglet.2014.12.001.

[10] W. Yu, T. Dillon, F. Mostafa, W. Rahayu, and Y. Liu, « Implementation of Industrial Cyber Physical System: Challenges and Solutions », *IEEE International Conference on Industrial Cyber Physical Systems (ICPS)*, Taiwan, 2019, pp. 173-178. doi: 10.1109/ICPHYS.2019.8780271.

[11] R. Luis de Moura, T. Monteiro Brasil, L. de Landa Ceotto, A. Gonzalez, L. Paulo Barreto, and L. Bassini Werner, « Industrial Internet of Things: Device Management Architecture Proposal », *International Conference on Computational Science and Computational Intelligence (CSCI)*, USA, 2019, pp. 1174-1178. doi: 10.1109/CSCI49370.2019.00221.

[12] Y. Zhang, C. Qian, J. Lv, and Y. Liu, « Agent and Cyber-Physical System Based Self-Organizing and Self-Adaptive Intelligent Shopfloor », *IEEE Trans. Ind. Inform.*, vol. 13, No. 2, 2017, pp. 737-747, doi: 10.1109/TII.2016.2618892.

[13] D. Wu, S. Liu, L. Zhang, J. Terpenny, R.X. Gao, T. Kurfessd, and J.A. Guzzo, « A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing », *J. Manuf. Syst.*, vol. 43, 2017, pp. 25-34, doi: 10.1016/j.jmsy.2017.02.011.

[14] J. Wang, P. Zheng, Y. Lv, J. Bao, and J. Zhang, « Fog-IBDIS: Industrial Big Data Integration and Sharing with Fog Computing for Manufacturing Systems », *Engineering*, vol. 5, No. 4, 2019, pp. 662-670, doi: 10.1016/j.eng.2018.12.013.

[15] H. Chegini, R. K. Naha, A. Mahanti, and P. Thulasiraman, « Process Automation in an IoT–Fog–Cloud Ecosystem: A Survey and Taxonomy », *IoT*, vol. 2, No. 1, 2021, pp. 92-118, doi: 10.3390/iot2010006.

[16] S. Hoppe, « OPC Unified Architecture Interoperability for Industrie 4.0 and the Internet of Things », 2020.

[17] P. P. Khine, and Z. S. Wang, « Data lake: a new ideology in big data era », *ITM Web Conf.*, vol. 17, 2018, pp. 03025, doi: 10.1051/itmconf/20181703025.

	Open source	EDGE computing	Cloud computing	Security	Interoperability	BIG Data	Machine Learning	Rules-based expert system
[7]	✓	✗	✓	✓	✗	✓	✓	✗
[8]	✓	✗	✗	✗	✓	✓	✓	✗
[9]	✗	✗	✗	✓	✗	✓	✗	✗
[10]	✓	✗	✓	✓	✓	✓	✓	✗
[11]	✓	✓	✓	✓	✓	✓	✓	✗
[13]	✓	✓	✓	✓	✓	✓	✓	✗
[14]	✓	✓	✓	✓	✗	✓	✓	✓
Our Proposed Achitecture	✓	✓	✓	✓	✓	✓	✓	✓

Table 1: Architectures Comparison

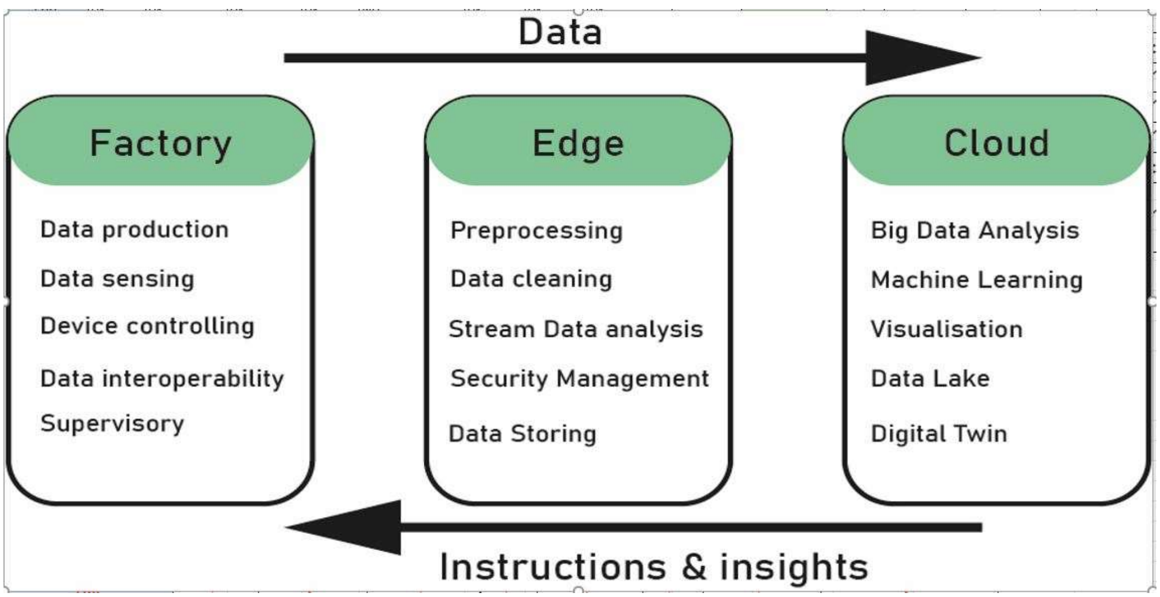

Figure 1: IIoT Data Flow

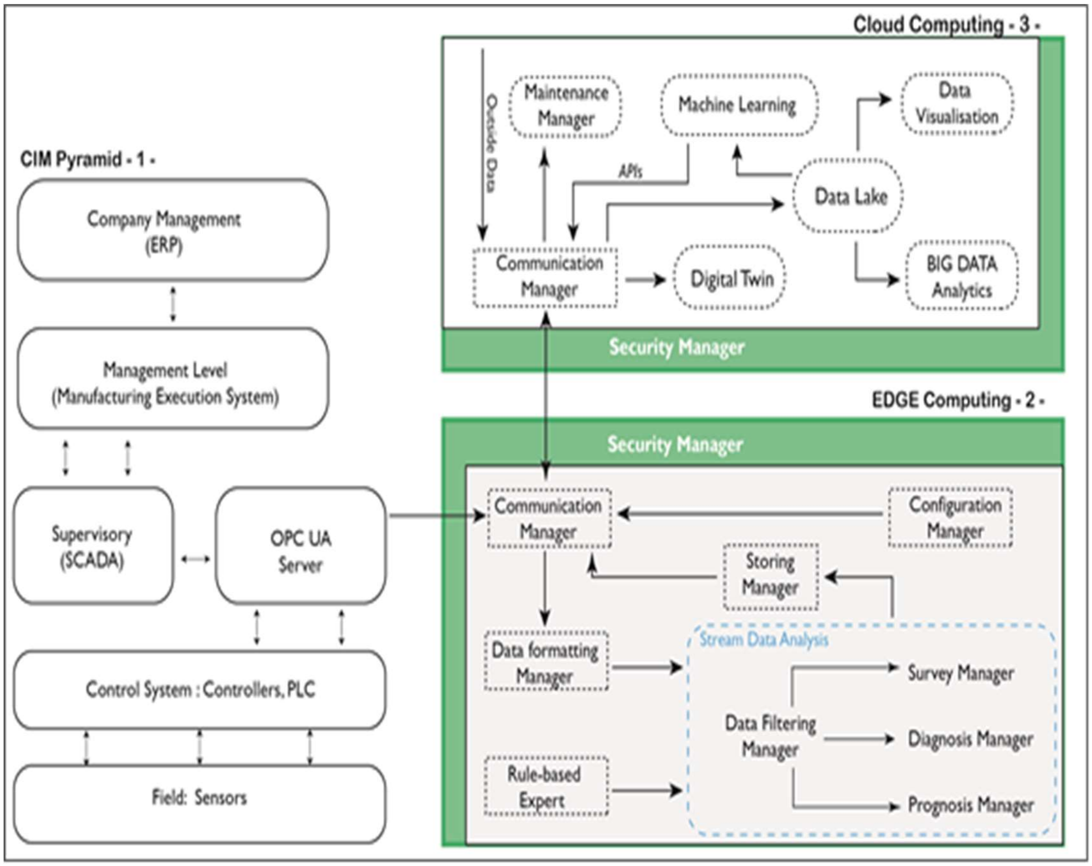

Figure 2: Proposed architecture for gathering and processing data