IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 25 February 2023, accepted 16 March 2023, date of publication 19 April 2023, date of current version 4 May 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3268223

RESEARCH ARTICLE

Strategies for Assuring Low Latency, Scalability and Interoperability in Edge Computing and TSN Networks for Critical IIoT Services

DAVID L. GOMEZ, GERMÁN A. MONTOYA, CARLOS LOZANO-GARZON, (Senior Member, IEEE), AND YEZID DONOSO, (Senior Member, IEEE)
Department of Systems and Computing Engineering, Universidad de los Andes, Bogotá 111711, Colombia

Corresponding author: Germán A. Montoya (ga.montoya44@uniandes.edu.co)

This work was supported by the Universidad de Los Andes.

ABSTRACT In recent years, there has been an exponential growth of IoT (Internet of Things) devices in the market and increasingly, they are being part of critical mission functionalities for many industries. For this reason, processes associated with these devices are generating new requirements at the level of interconnectivity, security, and scalability. Due to the large volume of devices, this growth process has tested the existing communications infrastructure, deployment architecture, different support services, technical requirements and the intrinsic characteristics of the services offered. In this sense, the use of Edge Computing schemes and TSN (Time-Sensitive Networking), is proposed as a tool for supporting IIoT (Industrial Internet of Things) services to achieve optimization and assurance processes of latency, scalability and interoperability parameters required by critical mission services such as health, transportation and production. In this sense, we propose several strategies for assuring low latency, scalability and interoperability in Edge Computing and TSN networks for critical IIoT services. These strategies are evaluated considering different levels of congestion and considering different schemes such as best effort, TSN and Edge Computing.

INDEX TERMS IIoT, TSN, edge computing.

I. INTRODUCTION

In recent decades, technology has gone from being an enabling element for the productive processes of society to being the dominant element in many aspects of daily life, transforming itself to be an active part of the global economy and guiding many of the trends of transformation of society [1].

The evolution of technology, especially the Internet, has led to the creation of the concept of a digital economy, where all interactions are carried out through a connected technological device (Internet - Intranet), accelerating the exchange of information and facilitating the process of exchanging goods, services and money [1].

As can be seen in Figure 1, technological processes have permeated multiple layers in different productive sectors, transforming the operations of each of them at different levels. This transformation has been achieved through much faster interaction processes such as e-commerce models

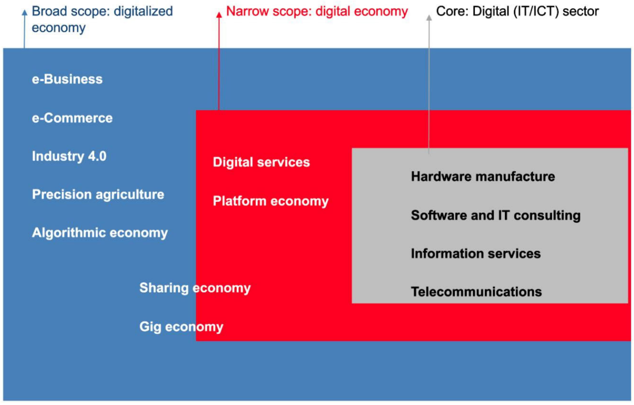

FIGURE 1. A representation of the digital economy [1].

The associate editor coordinating the review of this manuscript and approving it for publication was Theofanis P. Raptis.

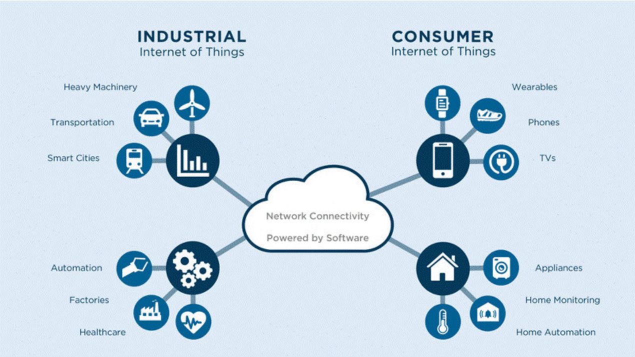

FIGURE 2. **Why IIoT is diferent from IoT [2].**

compared to the traditional model of bricks and mortar, going through processes of optimization of the agricultural process, allowing much more precise activities in the different stages such as sowing control, application of fertilizers, irrigation, and pest control until reaching what is known as industry 4.0. In industry 4.0, the productive operations of different sectors rely on information technologies to strengthen control and monitoring processes, reducing maintenance times, accelerating reaction times to events and reducing downtime, configuration and planning [1].

The technology responsible for supporting the aforementioned benefits is known as IIoT (Industrial Internet of Things), which is an evolution of the concept of IoT (Internet of Things). IIoT refers to a network of connected devices in the industrial sector that transfer data without human-to-human or human-to-computer interaction. IIoT creates an AI-powered "system of systems" that can curate, manage, and analyze data from one end of the business to the other. Within this system, machines, people, and other systems can work together in real-time powering more resilient operations and sustainable business growth [2].

Starting from the satisfactory results of the IoT model and the potential of communications between devices to accelerate the delivery of services, exchange of information and knowledge generated by the availability of additional information, the idea was born to apply a similar model for critical services within different industries such as the medical sector to strengthen patient monitoring processes through interconnected devices, manufacturing industries to strengthen control and monitoring processes of their operations, or manufacturing industries in automation processes of activities. Additionally, a new field of application was created: Smart cities and utilities where the new information generated by interconnected devices allows much faster decision-making to address any particular situation, reducing its impact or allowing the generation of considerably more precise action [2].

Although the general concept of IoT and IIoT is similar - interconnection of devices - the specific needs and requirements of each of the segments are very different. As can be seen in Figure 2, the services covered are totally different since, for example, availability requirements vary compared to a physical activity registration service through a smartwatch compared to production process control services of a factory. Therefore, approaches to support each of these branches tend to be increasingly different [2].

The central focus of this paper will be IIoT technologies and how new technologies, specifically advances in network platforms, can help meet the needs of these sectors.

The main contributions of this paper can be listed as follows:

- Currently, in terms of Industrial Ethernet behavior, the most used technological solution in the market is Industrial Ethernet, whereby a general model is proposed using the capabilities of this technology. The values obtained and validated with current industry processes are compared against new technologies and deployment strategies.
- With respect to TSN, basic requirements of TSN services are established with the aim of analyzing the network behavior, considering the impact of the algorithms or flow control models that TSN technology presents. The basic TSN services considered are Clock Synchronization (AS), Device Configuration (QCP), Flow reservation (QAT and QCC), and Route control (QCA). These services are described in detail in the proposal section.
- In addition, the behavior of TSN networks within an IIoT environment is approached by implementing standard topologies and messaging processes for IIoT platforms to establish a technology baseline. In this sense, two scenarios are proposed to be used: the use of traffic modelers based on credits (Qav) and time guards and packet breaking technologies (Qbv-Qbu). More details are described in the proposal section.
  Taking into account the previous considerations, we offer a final result consisting of the comparison of technologies used for modeling the services, which are evaluated considering latency, channel efficiency, jitter and queuing behavior. More details are described in the proposal section.
- With respect to Edge Computing technology, this technology has been implemented for structuring the network topology and services that can be deployed. Since computing nodes are distributed in multiple locations, optimizations in the general behavior of services can be achieved. In this sense, two scenarios are configured to confirm the advantages of this technology in terms of network performance and control: a centralized scheme and an Edge Scheme. More details are described in the proposal section.

The remainder of this paper is organized as follows. The general problem statement and the background are described in Sections II and II-A. In Sections III and IV, our proposal and its implementation are shown, respectively. In section V, we presents the results obtained, and its

TABLE 1. Network requirements for industrial services [3].

Economic Sector	Service Application	Latency	Latency fluctuation
Medicine	tele-surgery	3 – 10 ms	< 2 ms
Industry	Automatized industry Control systems	0,2 μ s – 0,5 μ s 25 μ s – 2 ms	N/A some μ s
Banking	High frequency transactions	< 1 ms	Some μ s
Aerial	AFDX	1 - 128 ms	Some μ s
Automotive	Driving assistance Power train motor Traffic and security	100 - 250 μ s < 10 ms < 5 ms	some μ s some μ s some μ s
Digital entertainment	Augmented reality Professional Audio/video	7 – 20 ms 2 – 50 ms	some μ s < 100ms

TABLE 2. Network requirements for massive consume [4].

Application Type			Tx Rate (Kbps)	Loss (%)	Latency (ms)
Data		FTP	Limitless	0	Not apply
Real time	Audio	Voice	≤ 64	10^{-4}	≤ 300
Real time	Audio	VoIP	10 to 64	$5 * 10^{-2}$	≤ 300
Real time	Video	MPEG4	≤ 2000	10^{-2}	≤ 40
Real time	Video	H.320	≤ 64	10^{-4}	≤ 40
Not Real time	Audio	CD	150	10^{-4}	Buffer size
Not Real time	Video	MPEG4	No limit	0	Buffer size
Games	Multi-protocol	Strategy games	20	$0 - 10^{-2}$	500
Games	Multi-protocol	First Person Shooter	20	$0 - 10^{-4}$	100
Network services		NFS	No limit	0	Not apply

discussion in Section VI. Finally, in Section VII, we shows the main conclusions of our work.

II. PROBLEM STATEMENT

IIoT platforms have had great growth in recent years thanks to their impact on the production processes of different industries, facilitating process configuration processes, improving response times, strengthening monitoring and control processes and, recently, reducing dependence on on-site personnel to control production processes. These elements have had great importance in the last year due to pandemic conditions that have restricted the movement of personnel and hindered the productive tasks of many industries [2].

The trend of automating production processes is not something new. Since the early 2000s, this type of technology has begun to generate contributions in different industry sectors. In recent years, the growth of productive processes has increased dramatically, where it is estimated that the IoT market will be around 1,500 billion dollars per year by 2025 [1].

With these accelerated expansion processes, new challenges for this technology are identified: from assurance processes to scalability and resilience models. Thus, migrating platforms that are much more sensitive to this type of technology require more requirements in comparison with current capabilities [2]. In this sense, based on these requirements, the current research is focused on Security, Availability, Interoperability, Scalability and Performance.

In this paper, we will focus on the elements of availability, scalability and performance of network platforms that support IIoT technology, which impose much stricter requirements on interconnection platforms in comparison with other similar interconnection services such as IoT schemes for end consumers. Connectivity parameters for different IIoT services are presented in Table 1.

The requirements of the most common services on the Internet are shown in Table 2. As can be seen, the latency and jitter requirements of IIoT services are much more restrictive than the requirements that a standard Internet service may pose. This is why the approach to network services that must be supported by Internet platforms IIoT must be much more robust and have sufficient technological capabilities to meet these constraints.

In many scenarios, the development of specialized solutions for certain market niches obtains the expected results. However, in the case of IIoT, many proposed solutions have generated a large number of standards in the sector, hindering interoperability processes between different solutions and generating very strong lock-in processes. For this reason, in order to satisfy the network requirements of IIoT services, it is necessary to seek the application of industry standards that allow the network solution architecture to be interoperable and provide rapid and easy growth in case it is required, enhancing industry flexibility.

IIoT models have managed to enter critical services of society to support their performance, optimization and data collection. However, by their nature, these services require very different operating parameters than personal/home IoT models. For this reason, it is necessary to propose more robust architectures and deployment schemes to meet the new requirements for critical/industrial services.

As seen in Figure 3, the main elements of society are undergoing a transformation process by making use of new IoT technologies, generating multiple deployment domains that allow for more fluid interaction. For example, by integrating domains of smart cities with supply chains, distribution routes can be optimized and the impact on vehicle traffic can be reduced. Another example is the integration of smart home services with smart industry services since it accelerates the information flow related to the use of products and consumption habits. All these possibilities open up a new world for industries, users and cities to build new services or optimize existing ones.

FIGURE 3. **IoT smart environment [5].**

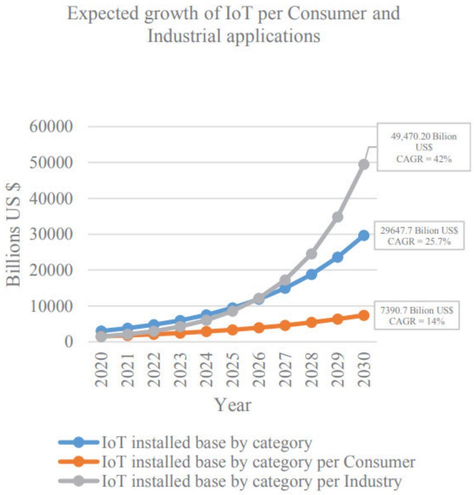

FIGURE 4. **Expected growth of IoT per consumer and industrial applications [5].**

IIoT models are one of the pillars of the fourth industrial revolution by enabling intelligent production models, allowing industries to respond more quickly and efficiently to market needs. In this sense, these devices will have more presence in our society not only as devices for the home but also as enablers of optimization and innovation processes within critical services.

According to projections for IIoT services, an accelerated growth in the market for this technology is observed (Figure 4), where industrial services will have the greatest growth in the next decade leveraged by industrial automation processes.

A. BACKGROUND AND RELATED WORKS

The theoretical framework is divided into 3 main elements:

- TSN (Time-Sensitive Networking): one of the main components of the proposed solution in order to ensure the network parameters required by critical services.
- Edge Computing: The second critical element of the solution is the integration of services under the Edge scheme. This component provides a more efficient flow of information by reducing access times to the required processing layers.
- 5G: Fifth-generation networks are becoming the standard for last-mile communications for IIoT devices thanks to their high speeds, coverage, and ease of deployment. This implies an additional challenge in the integration processes to ensure the interoperability of QoS levels between multiple technology domains.

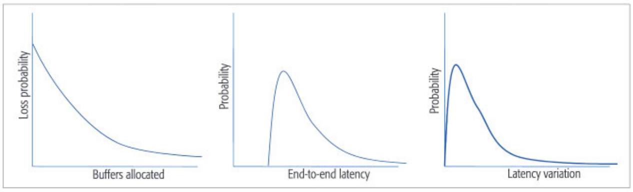

FIGURE 5. **Best effort [7].**

1) TSN (TIME-SENSITIVE NETWORKING)

The TSN technology is a standard defined by the IEEE that aims to provide a reference framework for service parameter assurance schemes in a ‘‘deterministic’’ way, using the existing infrastructure and managing to maintain coexistence between current transmission standards [6]. This technology offers schemes to be defined in which the maximum values of latency, the probability of loss or the jitter define maximum levels that can be aligned according to the requirements of a certain service. This is an advantage over the best effort model where there are some average values but there is no way to ensure that any packets do not comply with this expected behavior. Nowadays, there are models or specialized digital networks that ensure these parameters (latency, packet loss and jitter), but involve high costs such as CBR (constant bit rate packet) in which resources are provisioned throughout the entire network regardless of whether they are being used or not [7].

Figures 5, 6 and 7 present a comparison of the top 3 models: best effort, CBR and TSN. The best effort model maintains the service parameters in an expected range; however, by the intrinsic structure of the service, it is observed that there may be cases where these parameters may vary considerably. In the case of the CBR model, it is observed that the service parameters have a minimum variation achieved by ensuring resources throughout the network, which implies that the cost is much higher depending on the topology used. Finally, in TSN it is observed that the parameters are not as strict as in the case of CBR, but a substantial improvement compared to the best effort model is observed. This is obtained at a much lower cost than what alternative schemes can offer us, and most importantly, the maximum variation values are limited whereby these parameters can be guaranteed to services that are running on the network, which is not possible in the best effort model.

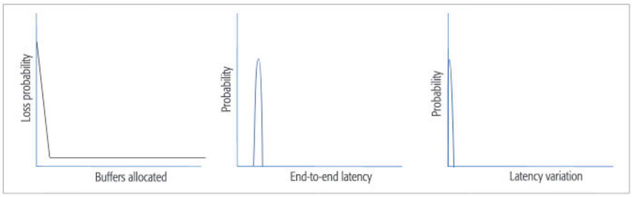

FIGURE 6. CBR [7].

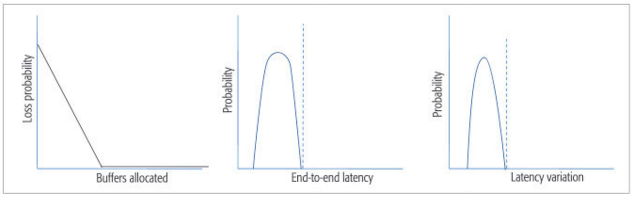

FIGURE 7. TSN [7].

What we currently know as TSN was born from an initiative known as audio-video bridge, in which operating schemes were proposed for real-time multimedia services for offering sufficiently robust network parameters in terms of latency, jitter or packet loss to achieve an optimal experience for the end-user. With the passing of time, this model was extended to any type of traffic that required such levels of service in the network, and the following standards that define TSN were proposed [8]:

- IEEE 802.1AS - Clock Synchronization
- IEEE 802.1Q - Bridged Networks
- IEEE 802.1BA - Audio Video Bridging
- IEEE 802.1Qbv, Qbu, Qci, Qch, Qcr - Queuing and Forwarding
- IEEE 802.1CB - Traffic Replication
- IEEE 802.1Qca, Qcc, CS - Resource reservation
- IEEE 802.1CM - Time Sensitive Networks for Fronthaul

Considering the exposed standards, TSN models a network with the following key elements [7]:

- Synchronization: In order to implement the necessary controls, TSN requires synchronization between all network actors to define transmission time allocation processes with the aim of ensuring appropriate queuing and latency levels expected by transmitters.
- Contracts: The most important aspect within TSN is to define certain levels of service required by applications in order to adjust the network to traffic requirements. These contracts must provide the inclusion of restrictions in terms of latency, jitter, and packet loss. At the same time, they must include or modify existing service contracts in order to optimize utilization of resources available in the network.
- Interoperability: Finally, interoperability must exist in TSN networks to include best effort models. In this way, available resources can be used for traffic that does not require such strict operating conditions and improve the network cost model.

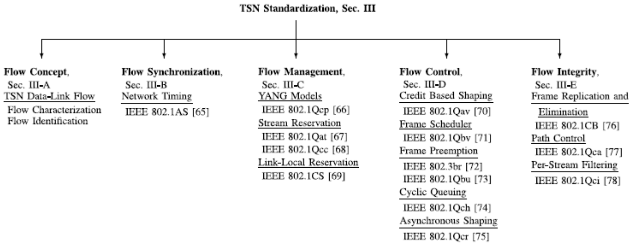

FIGURE 8. TSN standardization.

The TSN standards (Figure 8), although based on existing models since 2009, have only been formally released since 2015. Since then, multiple updates have been generated specifically in the queuing and allocation models of network resources.

Currently, the released TSN specification covers multiple fronts to achieve a set of services robust enough to provide all the tools for critical traffic modeling on Ethernet networks, fulfilling the established requirements for each traffic flow.

Current works on TSN networks are focused on traffic planning and queuing algorithms in order to optimize the use of resources that the network offers, whereby there are multiple standards for managing queues in TSN models. In this sense, multiple works focused on the use of these standards have emerged to adapt the network as best as possible to traffic conditions, ensuring compliance with the defined service contracts and, at the same time, optimizing the resources used in the network.

Among the works found, it is worth highlighting the work [13], which proposes the integration of TSN networks for IIoT services, defining real-time configuration heuristics for the optimization of traffic generated by sensors and actuators within an IoT model industry. In addition, in [24], the authors propose an efficient communication prototype that uses time-sensitive networking (TSN) and edge computing to reduce latency with zero-loss redundancy protocols that ensure the sustainability of IIoT networks with smooth recovery in case of unplanned outages. However, this work, in comparison with our proposal, only refers to analyze latency and lacks analyzing queue behavior and jitter in these kinds of networks.

2) EDGE COMPUTING

With the consolidation of cloud services, many operations locally performed in data centers have been migrated to cloud services in order to manage scalability and redundancy schemes with lower costs compared to implementations of the same services locally. However, with the inclusion of new technologies and the management of interconnected environments, new requirements have appeared to test the cloud-based operation scheme:

- Latency: with the inclusion of critical systems in interconnected models, latency came to play a critical role for many systems since, due to network schemes and

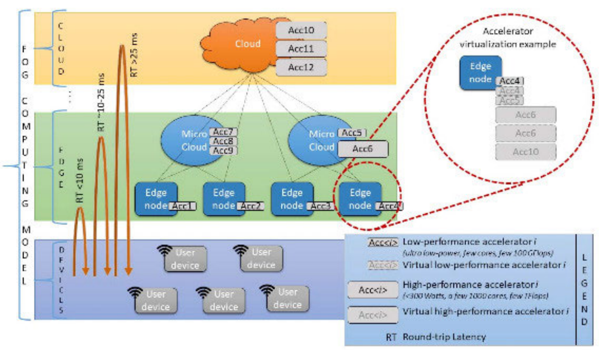

FIGURE 9. EDGE computing model [11].

computing infrastructure, the response times obtained for specific services surpassed acceptable ranges.

- Volume of operations: the volume of information generated has grown exponentially, stressing communication networks and computing infrastructure. In addition to this, not all the information generated is required for online decision-making processes or may not even be valuable for the provided service.

One of the options to solve these problems consists of distributing operations between different layers, reducing latency by placing processing nodes closer to the final devices in order to make decisions with time restrictions and, at the same time, filtering the consumption of services performed on the cloud.

Within the analyses to be carried out for the implementation of an EDGE model (Figure 9), in [11] and [12], the authors present the following advantages and challenges:

- Capabilities: it is important to consider the available capabilities for the services deployed in the EDGE with the aim of avoiding overuse or underuse of these services.
- Test environment: when moving to the final devices to add computing and storage capabilities, the heterogeneity of the devices and the lack of a valid test environment for their execution are found.
- Neutral Systems: similar to the previous case, the end devices have evolved by manufacturers' policies, whereby it is necessary to achieve interoperability between different systems to have a successful deployment of the EDGE model.
- Coordination: coordination between different applications in the EDGE nodes implies a process of coordination between them in order to attend complex operations, which can be a titanic task depending on the devices we want to cover.
- Privacy and Security: by placing processing nodes outside the cloud scheme, it is necessary to analyze and implement applicable security requirements for deployed services, considering capabilities of used devices.

On the other hand, the opportunities are the following:

- Bandwidth consumption: by providing operations to be carried out on edge devices or near them, decisions can be made for sending only critical data or operations to the cloud, achieving a reduction in bandwidth consumption in the cloud.
- Latency reduction: one of the basic points of an EDGE model is the reduction of application latency by processing nodes close to edge devices in order to process information and make decisions about these applications.
- Communication parameters: by having processing nodes closer to devices that generate information, it is possible to define schemes to ensure communication parameters in a simpler way compared to a pure cloud scheme, where information would need to travel through multiple providers to reach the destination.

Taking into account the EDGE model, there are multiple research studies focused on evaluating the feasibility of an EDGE scheme for specific services. In the context of this research work, there are multiple approaches to enable IIoT service optimization processes through the use of EDGE computing schemes. Most worked topics are the technology integration processes to overcome the limitations caused by the heterogeneity of edge devices, frameworks for providing the integration and distribution of operations between EDGE nodes [16]. In general, research studies are focused on determining the viability of distributed models under EDGE technologies and what benefits they can bring to the underlying services to optimize the proposed requirements [13], [17]. In addition, in [25], the authors propose a three-layer task offloading framework at which tasks with high computing requirements are offloaded to the cloudlet layer and cloud layer. Tasks with low computing and high communication costs are executed on the device layer, whereby their proposal avoids transmitting large amounts of data to the cloud and can effectively reduce processing delay. However, this work, in comparison with our proposal, only refers to analyzing delay and lacks analyzing queue behavior and jitter in these types of networks.

3) INDUSTRY 4.0

With the arrival of new devices for providing information exchange, production processes have undergone a drastic change in their operation scheme since the speed of information flow has increased exponentially. This has allowed decision-making with information in real-time, and at the same time, it has allowed changes in the operation schemes of previously disconnected devices to be carried out immediately to adapt to the new market needs.

These operational changes propose a paradigm shift, where products and production processes are smart and react to the changes required by the market and the industry. This new scheme defines a set of very specific requirements on communication technologies which have been covered with specialized technologies in the industry such as IRT or EtherCAT [9]. However, with the massification of information technologies in society, new migration processes of technologies from the final consumer to the industry using standard market technologies have appeared.

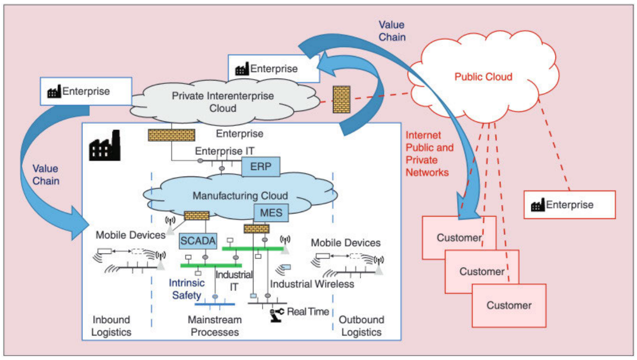

FIGURE 10. **Industrial automation schemes [9].**

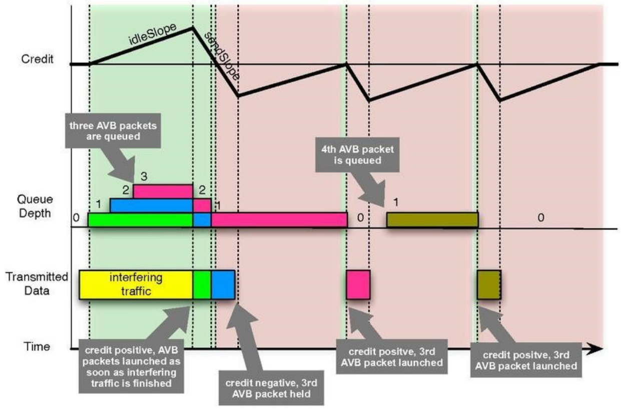

FIGURE 11. **QAV Credit Base Shappers [18], [19].**

As can be seen in Figure 10, by transferring technology from the final consumer to the industry, new challenges must be addressed, such as security issues, scalability and response times, which currently are research topics to optimize deployment schemes.

To encapsulate all the changes originated from the inclusion of new technologies in the industry is complex, but Lorenzo Bassi carries out a work in which he establishes a set of characteristics that define Industry 4.0, establishing that indeed it is a paradigm shift. These set of characteristics are: intensive Internet usage, Flexibility, communication, virtualization, and cyber-physical systems (CPS) [10].

All these elements are combined using technologies for providing fast information exchange (IoT), applied to the industry (IIoT), achieving a production model handled by the collected real-time information.

The application of new technologies in the industry has brought new challenges which are currently the subject of research, among which we found the following main topics: Communication, Security, and Data Processing [14]. Among the works found, it is worth highlighting the work [26], in which the authors propose a new architecture that attempts to satisfy several requirements such as scalability, heterogeneity of information across different applications, and efficient aggregating operation. However, this work, in comparison with our proposal, implements an architecture but lacks evaluating scenarios for obtaining results in terms of latency, jitter, and queue behavior.

III. PROPOSAL

This work hopes to determine the following elements within an IIoT services framework using the latest available technologies:

- Industrial Ethernet behavior: Currently, the most used technological solution in the market is Industrial Ethernet, whereby a general model is proposed using the capabilities of this technology. The values obtained and validated with current industry processes will be compared against new technologies and deployment strategies. The specialized protocols that will be considered are: Modbus, Fieldbus, EtherCAT, Profinet, and EtherNet/IP.
- TSN Basic Elements: Basic requirements of TSN services will be established with the aim of analyzing the network behavior, considering the impact of the algorithms or flow control models that TSN technology presents. The basic services proposed are:
  - Clock Synchronization (AS): By having topologies with clock synchronization processes, it is possible to use time-based transmission models or by controlling traffic through time lapses.
  - Device Configuration (QCP): A dynamic topology configuration generates simpler increasing or decreasing (failures) processes and reduces the process of exchanging messages to establish the real state of the network.
  - Flow reservation (QAT and QCC): One of the main elements within network control processes for TSN is the ability to reserve flows for different applications. For this reason, reservation and control models are proposed in TSN to have a framework for flow control algorithms and protocols.
  - Route control (QCA): Although this process may have some modifications, it is proposed to use the IEEE standard defined for controlling routes within topology (QCA) and in this sense focus on control flow protocols and algorithms on established routes by this standard.
- TSN behavior: The behavior of TSN networks within an IIoT environment will be determined, starting from standard topologies and messaging processes for IIoT platforms in order to establish a technology baseline. For this stage, two scenarios are proposed to be used:
  - Qav: In the first stage, the use of traffic modelers based on credits is proposed for determining the behavior of the network without the use of time multiplexing in the channels (see Figure 11).
  - o Qbv-Qbu: the second stage uses time guards and packet breaking technologies (packet preemption)

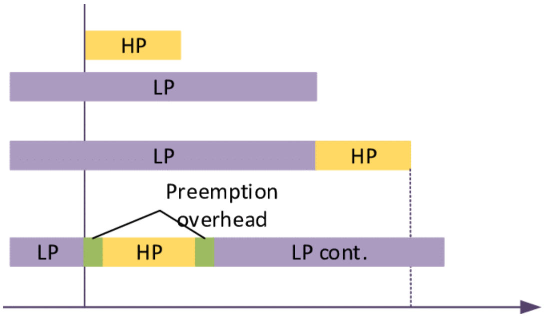

FIGURE 12. **Packet preemption [20].**

to ensure network parameters in critical flows (see Figure 12).

- Comparison of results: with the different stages carried out, the final result is the comparison of the different technologies used for modeling the services, which will be evaluated in the following aspects:
  - Latency: Latency management of different flows transmitted in the topology by establishing the level of fulfillment of the proposed requirements and the behavior of this value at different levels of resource saturation.
  - Channel efficiency: This process determines the real efficiency of the use of the network infrastructure, considering how many flows can be transmitted within the proposed topology, and what is the underutilization of resources associated with the constraints imposed by each flow control model of the associated technologies.
- Edge Computing: With the management of critical services with such strong network restrictions, the controls are not only applicable at the network level. The deployment of computing capabilities has a fundamental role in achieving the desired behavior for industrial services. For this reason, Edge computing offers alternatives for structuring the network topology and services that can be deployed. Since having computing nodes distributed in multiple locations, optimizations in the general behavior of services can be achieved. In this sense, two scenarios will be carried out to determine the possible advantages of these schemes in terms of network performance and control.
  - Centralized scheme: The base model is a centralized standard deployment, where all the computing resources are deployed in a central node (local data center), and the different elements of the network (sensors, actuators, etc.) send the traffic through the network infrastructure (Industrial Ethernet or TSN).
  - Edge scheme: In this scheme, the capacities are distributed in different topology locations, allowing multiple processing nodes and reducing the agglomeration of traffic towards a single network point.

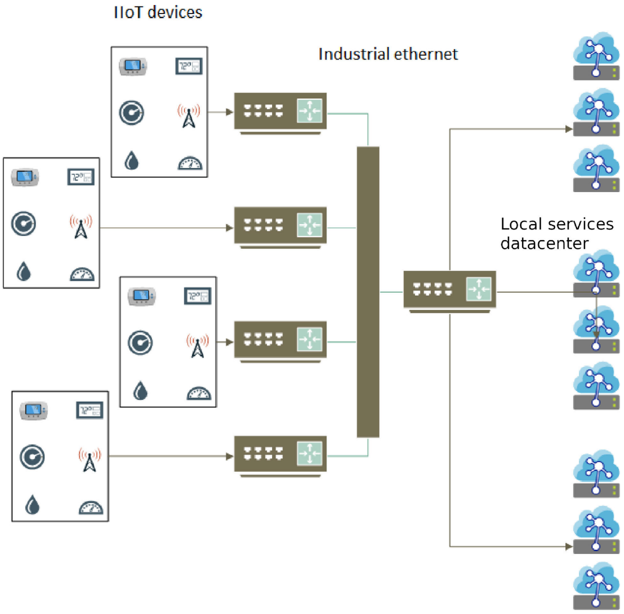

FIGURE 13. **Industrial ethernet default scenario.**

These two models will be evaluated by determining the effect they can have on the behavior of the network parameters (Delay, and Channel Efficiency) for the different flows that are transmitted in the infrastructure.

A. SCENARIOS

The following general scenarios have been defined in order to obtain base data and validate the implementations of new technologies in the context of IIoT services. Within each scenario, three levels of traffic concurrence will be handled in order to establish the behavior of the main study indicators (latency, jitter and channel efficiency).

- Low consumption: At this level, the infrastructure has a maximum utilization of 20% of the total capacity, measured in the bandwidth utilization values at the point of the topology with the highest levels.
- Medium Consumption: This level represents a utilization of 50% of the topology's capacities, measured in the bandwidth utilization values at the point of the topology with the highest consumption.
- High Consumption: This level represents a utilization of 80% of the measured resources, at the point of the topology with the highest bandwidth usage.

1) BASIC SCENARIO

In Figure 13, a first scenario implements current technologies to achieve the comparison base data. In this case, Industrial Ethernet will be used to model all network connections between IIoT devices and control services that are locally deployed. For this scenario, the standard scheme for centralized deployment of control services for different IIoT services is proposed.

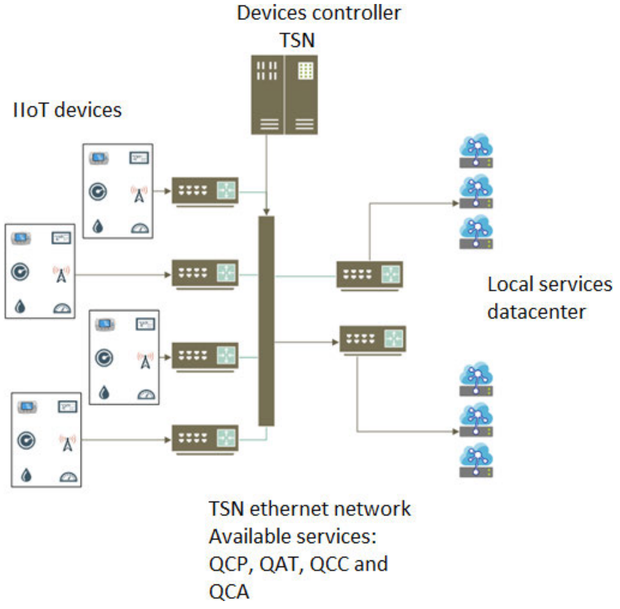

FIGURE 14. **TSN scenario - default services.**

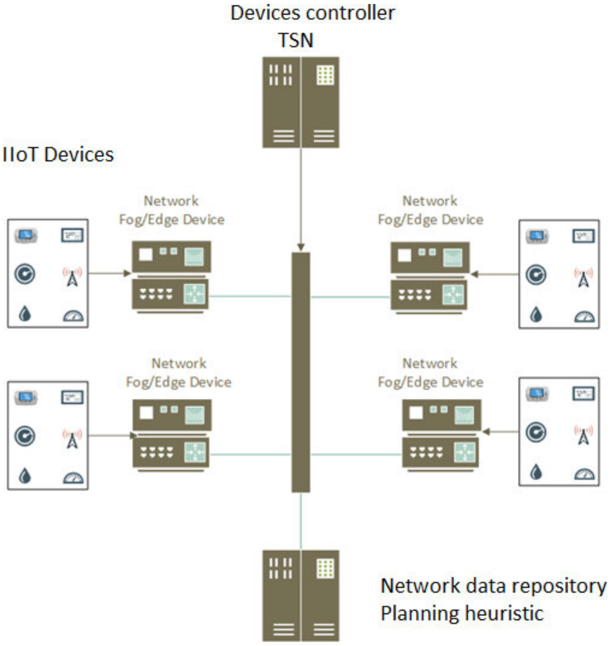

FIGURE 15. **TSN scenario.**

As a result of this scenario, comparison values will be obtained for the different technologies, which will be evaluated in the following scenarios using TSN networks and Edge computing schemes.

2) TSN

The second scenario (Figure 14) implies the implementation of TSN services within the topology, starting from the basic services required for its operation (QCP, QAT, QCC, QCA). No changes are made in the distribution of the deployment of the services, since it is necessary to determine the real impact of the TSN technologies in the main measurement indicators that have been defined.

Within this scenario, queuing and traffic management models are proposed within the network topology (QAV and Frame Preemtion), which will be simulated considering the three conditions of use that are defined for the scenarios. This scenario provides the necessary information for the first evaluation against the initial base scenario of Industrial Ethernet.

3) EDGE COMPUTING

In the context of this work, Edge computing technology allows us to adjust the service deployment scheme in order to facilitate the network requirements compliance for the different monitoring and control services for IIoT services. By having multiple nodes for the destination services, it is possible to balance the traffic in the topology, reducing the bottlenecks that could arise when adding multiple data flows to a single destination.

In this scenario (Figure 15), the TSN capabilities of the topology are maintained, but deployment models are unified so that using SDN (Software Defined Networks) and Edge nodes, multiple points of service attention for IIoT devices are modeled in the topology. For this scenario, the data collection services established in the previous scenario are maintained to identify the flows behavior, and thus, establish more efficient deployments of Edge nodes. It is clear that in terms of latency and network traffic, the option that gives better results is to deploy the services as close as possible; however, this solution is not always technically or economically viable due to many control software cannot operate in a effective way in distributed environments, and the costs associated with Edge nodes would increase the cost of implementing and maintaining the topology. The final result of this scenario is a heuristic that complements the one generated in the previous scenario, which can indicate possible changes to the topology indicating the inclusion of Edge nodes, taking into account the following premises:

- The aim is to reduce the capabilities of the network devices (Channels and Bandwidth) while maintaining compliance with the agreed requirements for each of the modeled priority flows.
- It is assumed that the different destination services (Control Services) do not operate in a distributed manner, whereby it is possible just one instance within the topology serving this traffic.
- To reduce the number of Edge nodes that are deployed in the topology. This allows us to control the complexity of the topology, and implementation and administration costs.

B. SCOPE

The proposal is defined to cover the following deployment elements within the IoT ecosystem:

- IoT networks - edge devices
- Edge device networks - Edge nodes

- Orchestration nodes
- Edge Node

Considering the defined scenarios, the solution proposal is aimed to strengthen the services deployed in Edge schemes, allowing us to control the information flow through TSN and performing processes through the optimization and deployment model to be built. Any operation outside these features would not be covered by the solution model. At the level of resources and deployment complexity, it is necessary to have connectivity and hardware capabilities that allow meeting the availability, latency and computing capacity requirements needed by the services, as well as are compatible with the standards to be used (TSN, Edge). As a final result, we propose a service deployment platform based on TSN networks with processing capacity in Edge nodes.

C. LIMITATIONS

The study assumes the availability of TSN services throughout the whole topology and that all devices are compatible with the latest standards defined by the IEEE.

All the traffic modeled within the topology is considered relevant, whereby it is necessary to generate configurations to accomplish the network constraints. Likewise, it is decided which non-relevant traffic will be modeled with best-effort schemes without affecting the behavior of priority traffic.

For IIoT devices connected by different technologies, such as 5G networks, the calculation of latencies and traffic restrictions is assumed until the point of traffic ingress to the TSN networks. Our proposal does not consider behavior modifications in current 5G networks.

Finally, our modifications are focused on strengthening the deployment platform. Our work never proposes optimization processes in terms of the internal operation of the services deployed on the platform.

IV. IMPLEMENTATION

In order to achieve results adjusted to reality, the Omnet ++ platform was selected, which allows us modeling discrete events quickly and manages multiple additional projects that facilitate the modeling of network topologies.

There are three main components within the simulation model, which will be detailed below:

- OMNET: this platform is the basis for the entire simulation process, since it exposes all the necessary components for the construction of the different required simulations. Within its general architecture, two critical elements are established. The .NED files, which are used to model the network to be tested. In these files, the main components of the topology and their different interactions are defined. The second element is the .INI files, which are used to establish the conditions of the simulation as well as the behavior of the elements defined within the topology. All the simulations carried out were worked under version 5.6.2 of Omnet ++. Omnet is an open source version designed for research and educational activities. It has commercial versions known as Omnest supported by Cogitative Software FZE. For more information go to https://omnetpp.org [21].
- INET: is a set of plugins built for Omnet, which allow the simulation of data networks by including multiple components that can be added in the construction of network topologies. These elements are designed to simulate real conditions taking into account all factors such as processing time, queuing, transmission time, etc. All the simulations carried out were worked under version 4.1.2 of INET. INET is a constantly evolving academic project developed by multiple educational and research institutions, as well as independent developers. For more information go to https://inet.omnetpp.org [21].
- Nesting: Nesting is an auxiliary project of an academic nature that extends part of the capabilities exposed by the INET components to include TSN functionalities in order to facilitate simulation processes. This tool includes clock synchronization handling, prioritized queuing, and packet breaking. For more information go to https://gitlab.com/ipvs/nesting [22].

A. PSEUDOCODE

In this section, a general pseudocode is provided to present the main features considered for evaluate the different technologies previously described.

Algorithm 1 Simulations Pseudocode.

```
Define General Simulation Parameters :
    simulation time = 300 secs
    MTU = 1500 Bytes
    MSS = 1492 Bytes
    processing time of Switches = 2000ns
    priority Queues supported = 8
    buffer size limit = 363360 Bits
Define Application Types
Select technology
Select consumption type
if technology = Industrial Ethernet then
    Define network connections
    Define control services
    Apply centralized services deployment
end if
if technology = TSN then
    Apply basic services
    Apply QAV queuing management
    Apply Frame Preemption

end if
if technology = Edge then
    Apply Traffic Balancing
end if
Evaluate Technology Performance
```

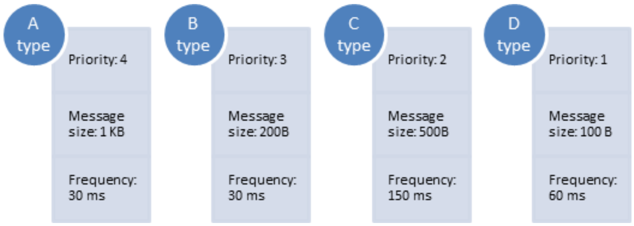

FIGURE 16. **IIoT applications type.**

In line 1, the general simulation parameters are established according to Section IV-G. Line 2 defines the application types to be used in the simulations, that is, applications A, B, C and D (See Section IV-B). Line 3 indicate the technology to be used in the simulation, that is, the Industrial Ethernet, TSN or Edge. In line 4, we indicate the consumption type to be tested: low, mid or high consumption. Lines 6 to 8, we establishes all IoT network connections, we deploy locally the control services, and a centralized deployment is applied for the control services. In line 11, the basic services of TSN are launched for its normal operation: QCP, QAT, QCC, QCA. In lines 12 and 13, the QAV queueing management and Frame Preemption is applied according to what was explained in Section III. In line 17, a traffic balancing is applied to reduce bottlenecks as several flows are added to a unique destination. Finally, in line 25, a performance evaluation is provided according to the technology previously selected.

B. APPLICATION CHARACTERIZATION

Within the simulation design, the first stage requires characterizing the type of applications that will generate traffic in the topology, for which the following conditions were established:

- Communication flow: It was established that the priority flow of information will be from IIoT devices to edge services, using a centralized scheme or using Edge computing models.
- The measured end-to-end latency is established as a performance metric. This value accumulates the times associated with transmission, queuing and processing.
- Although messaging schemes such as MQTT are used in some IIoT services, direct communication with the destination service is proposed, thus, eliminating the saturation effect that a messaging component established in the middle can generate [23], [24].
- The applications are characterized by periodic traffic of different sizes and with different frequencies in order to simulate the handling of different IIoT devices in the topology.

For all the simulations, we will work with the following types of applications defined in Figure 16.

To reduce the saturation effect that can occur when the simulation is activated, all the nodes that generate traffic will work with an exponential function distributed around 1 second.

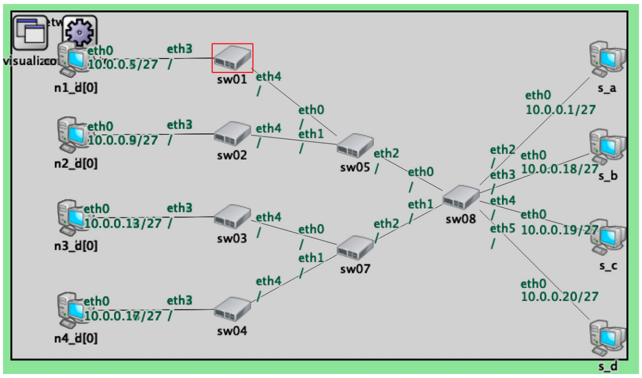

FIGURE 17. **Best effort and TSN topology.**

C. BEST EFFORT TOPOLOGY AND TSN

For the "Best Effort" and TSN scenarios, a centralized topology is proposed, as can be seen in Figure 17.

Within the proposed topology, there are the following base elements:

- Connection speed: In order to reduce the number of nodes for the simulation processes, all the connections observed in the topology work at speeds of 10 Mbps.
- Ethernet: All the topology works under Ethernet, regardless of whether it is for the "Best Effort" or TSN model.
- Distribution of nodes: All nodes (IIoT devices) are distributed evenly among the different edge switches, and the number of nodes may vary depending on the saturation scenario that is being evaluated.
- At the right part of the topology, the messaging destination services of the IIoT devices will be deployed with independent nodes (servers) to simulate different applications for monitoring different elements.

The topology architecture obeys a general network deployment scheme: edge switches are managed to connect the end devices; Intermediate aggregation switches for grouping the edge connections; and a final switch in the destination data center for supporting the connections of the servers that attend the different IIoT services.

D. EDGE LEVEL 1 TOPOLOGY

For the following scenarios, the previous topology was modified, generating multiple attention nodes for type D services. Take into account that an attention node in the context of our work and in general, in telecommunications networks, is a device or a point of access that is useful to control and monitor data traffic between two other points in the network. These points are indicated in multiple scenarios in our work. In addition, for us, it is an important node due to it monitors the quality of the transmission of data with the aim of improving the overall performance of the network. In our case, it was useful to evaluate latency, jitter and queue behavior in order to compare the different strategies proposed in our work.

As can be seen in Figure 18, the Edge capacity was modeled as the set of an attention node associated with Switch 05 and Switch 07. These nodes are connected to the network

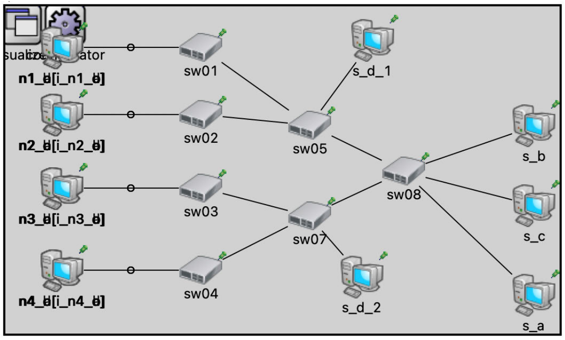

FIGURE 18. Edge level 1 topology.

by 10 GB links. For this reason, the transmission effect is negligible compared to the transmission rates of the rest of the 10 Mbps topology.

Within the topology, the following conditions are added to the existing conditions of the base topology:

- Independence of nodes (Services): It is assumed that for the management of type D services, the nodes sd1 and sd2 can operate independently and do not require traffic between them to attend to IIoT requests.
- Homogeneous Distribution: a homogeneous distribution of the attention of the requests is proposed. The IIoT devices of service D connected to sw01 and sw02 will be attended by the service deployed in the node sd1, and the IIoT devices of service D connected to sw03 and sw04 will be attended by the service deployed on node sd2.

In addition, the topology remains the same as the proposed in the previous scenario.

E. EDGE LEVEL 2 TOPOLOGY

For the following scenarios, the worked topology was modified, generating multiple attention nodes for type D services. As can be seen in Figure 19, the Edge capability is modeled as one node related to Switch1, Switch2, Switch03 and Switch04. These nodes are connected to the network by 10 GB links. For this reason, the transmission effect is negligible compared to the transmission rates of the rest of the 10 Mbps topology.

Within the topology, the following conditions are added to the existing conditions of the base topology:

- Independence of nodes (Services): It is assumed that for the management of type D services, the nodes sd1, sd2, sd3 and sd4 can operate independently and do not require traffic between them to attend to IIoT requests.
- Homogeneous Distribution: A homogeneous distribution of the attention of the requests is proposed. The IIoT devices of service D connected to sw01 will be attended by the service deployed in node sd1. The services deployed to sw02 will be attended by the service deployed in node sd2. The services deployed to sw03 will be attended by the service deployed to node sd3, and the services deployed to sw04 will be attended by the service deployed to node sd4.

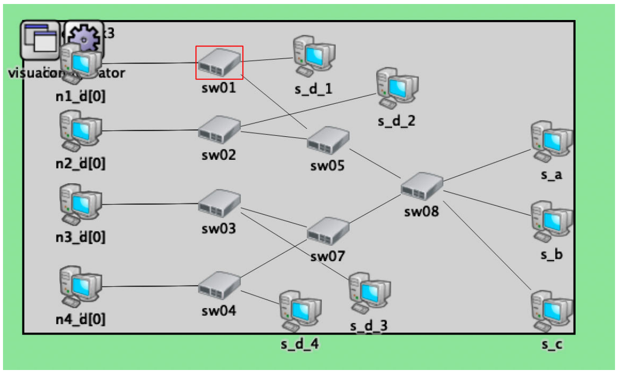

FIGURE 19. Edge level 2 topology.

TABLE 3. Bandwidth calculation.

Flow type	Packet size	P. time[1] (ms)	W. time[2] (ms)	R. lapse[3] (ms)	F. time[4] (ms)	BW[5] (bps)
A	1000	10	10	10	30	266.667
B	200	10	10	10	30	53.333
C	500	50	50	50	150	26.667
D	100	20	20	20	60	13.333
						Total BW: 360

[1] Processing time.
[2] Waiting time.
[3] Reconnection lapse.
[4] Frequency time.
[5] Bandwidth.

TABLE 4. Amount of IIoT required devices.

BW required	BW available	20% of Utilization	50% of Utilization	80% of Utilization
360000	10000000	5.555	13.888	22.222

Apart from the aforementioned conditions, the topology remains the same as the proposed in the previous scenario.

F. SATURATION SCENARIOS

For the generation of the saturation scenarios, we started from the characterization of applications that are going to work in the topology, and they were analyzed in the initial design (Figure 17). With this information, it was established that the critical points of the infrastructure are occurring in the connections from the aggregation Switches to the Services Switch (far right). In particular, the connections in question are those from Sw05 and Sw07 to Sw08, since these connections are supporting the traffic from two of the four edge switches that the topology handles, that is, half of the total traffic. In order to estimate the number of nodes to be used in each scenario, the calculation of the bandwidth required for the set of four applications is shown in Table 3.

With these values, we calculate the amount of IIoT required to reach the proposed saturation levels of 20%, 50% and 80%, as shown in table 4.

With the values obtained, an approximation to the following even number is made as follows: the bandwidth available to be used considering the BW available and the BW required corresponds to 10000000/360000, that is, 27.777. Then, for 20%, 50% and 80% of utilization, we calculate the number of devices, that is, $27.777 * 0.2$, $27.777 * 0.5$ and $27.777 * 0.8$, that is, 5.555, 13.888 and 22.222, respectively. These values were rounded up to the nearest integer, that is, 6, 14 and 23, respectively. Then, these values imply the amount of IIoT devices for two Switches at the end, that is, we duplicate the previous values, by means, 12, 28 and 46. In other words, 12 will be the amount of IIoT devices per application class when the percentage of utilization is 20% (Low congestion scenario); 28 will be the amount of IIoT devices per application class when the percentage of utilization is 50% (Mid congestion scenario); and, 46 will be the amount of IIoT devices per application class when the percentage of utilization is 80% (High congestion scenario). Finally, due to we are assuming four types of applications (services A, B, C and D), the total amount of IIoT devices for each congestion scenario are $12*4$, $28*4$ and $48*4$, that is, 48, 112 and 184, respectively. Therefore, the previous description can be summarized to generate the following scenarios:

- Low Congestion Scenario
  - IIoT devices amount per application class: 12
  - Total IIoT devices in topology: 48
- Medium Congestion Scenario
  - IIoT devices amount per application class: 28
  - Total IIoT devices in topology: 112
- High Congestion Scenario
  - IIoT devices amount per application class: 46
  - Total IIoT devices in topology: 184

G. GENERAL SIMULATION PARAMETERS

For the execution of the simulations, general parameters were established within the proposed topology:

- Simulation time: 300 seconds
- MTU: 1500 Bytes
- MSS: 1492 Bytes
- Switch Processing Time: 2000ns
- Supported Priority Queues: 8
- Maximum Buffer Size: 363360 bits (100 packets of 1500 bytes)

The general parameters of the simulation were selected in such a way that they allow adjusting to the behavior of real cases. For this reason, it was taken 1500 Bytes for the MTU, since it is the default value in most networks that are currently deployed, and therefore, the MSS value was left at 1492 Bytes, subtracting the TCP header. Regarding the capabilities of the Switches, a processing default value of 2000ns was established; however, this value does not vary for any of the services, whereby there is no impact on the statistical behavior of the simulations. In the queues, it was proposed that the entire topology is capable of handling the eight priority queues defined in the IEEE 802.1P standard. For TSN cases, the entire topology is capable of handling traffic prioritization models through credits and Frame Preemption processes. Finally, although within the scope a detailed behavior analysis of the queues within the switches was not proposed, a value of 100 packets of the maximum MTU was determined to avoid the discarding of packets. For the definition of the simulation time, previous analyses of the behaviors of different topologies were carried out, and it was found that most of them reached stable behaviors around 10 seconds. Therefore, the time was extended to ensure coverage of any atypical case and accomplish the expected saturation levels. With this, 300 seconds was established as the simulation time.

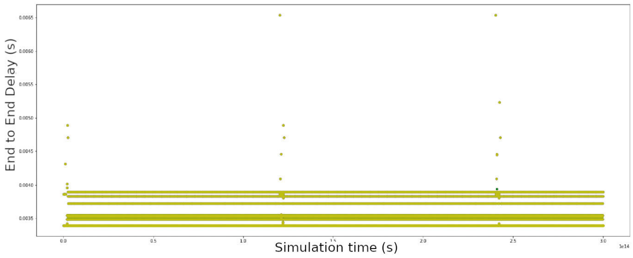

FIGURE 20. End to end delay vs simulation time - services a comparison low congestion.

All the information from the simulations was stored in vector files in SQLite format and processed with Python scripts for the generation of graphs and statistical analysis of the results.

V. RESULTS

Complementing the configuration specified in the Implementation section (IV), for sections V-A,V-A, and V-A (Low, Mid and High Congestion, respectively), it was performed a comparison between Best Effort and TSN, including two scenarios for optimizing the topology using EDGE schemes and using TSN capabilities. In addition, as we mentioned in the implementation section, the results were obtained considering three congestion levels. Finally, for obtaining these results, it was used Omnet 5.6, the package INET 4.4 with the Nesting complement [22]. More details are described as follows:

- Best Effort: basically all traffic has the same priority and use the same queue.
- General scheme for TSN:
  - Network synchronization: 802.1As
  - Flows filtering: 802.Qci
  - Time Aware Shaper: 802.1 Qbv
  - Gates configuration: 802.1Qcc

A. LOW CONGESTION

1) COMPARATIVE ANALYSIS OF SERVICES

Considering all the scenarios proposed, a comparative analysis is carried out for each service.

- Service A: Service A is the one with the lowest priority; however, due to it has a low saturation in the topology, no relevant changes are observed in the end-to-end latency behavior in the four scenarios as can be seen in Figure 20.

  Figure 20 shows the simulation results considering the service A for Best Effort, TSN, Edge optimization 1 and

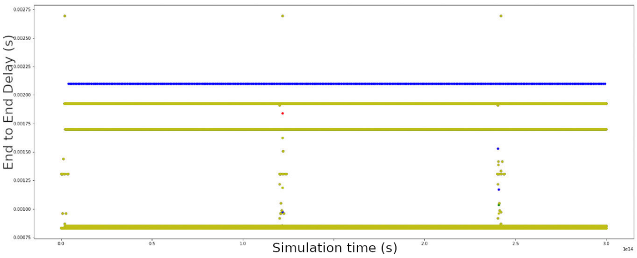

FIGURE 21. **End to end delay vs simulation time - Services B comparison low congestion.**

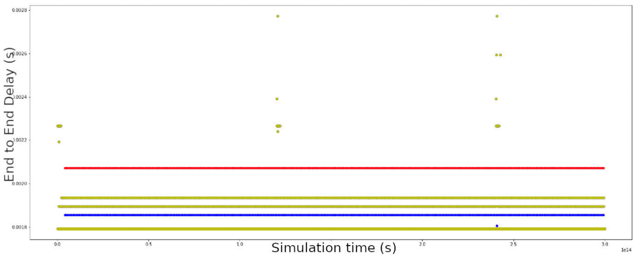

FIGURE 22. **End to end delay vs simulation time - Services C comparison low congestion.**

Edge optimization 2 technologies (red, blue, green and yellow, respectively).

- Service B: Service B has a higher priority than service A. Although there is an improvement in the last scenario (Edge Optimization 2), it is not too high because the topology is not saturated. The improvement presented is due to the reduction of traffic that circulates in the topology, since the traffic of service D is attended in the edge node at which the IIoTs are connected.

  Figure 21 shows the simulation results considering the service B for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively).
- Service C: In this service, behavior improvements are experimented by achieving an end-to-end latency reduction. This is due to two key points: the traffic prioritization and the reduction of the saturation of the network when attending the D service traffic in the EDGE nodes.

  Figure 22 shows the simulation results considering the service C for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively).
- Service D: This service shows improvements in its latency behavior, even with low levels of network utilization, since all the proposed optimizations benefit the service directly. As expected, the service using EDGE technology has extremely low latencies and low variability.

  Figure 23 shows the simulation results considering the service D for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively).

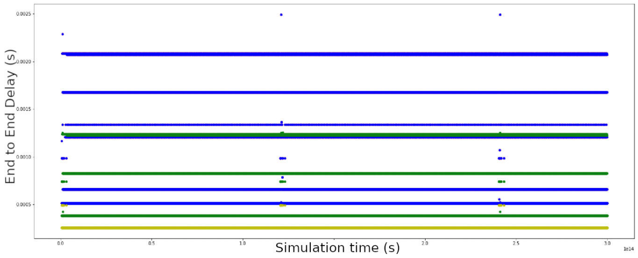

FIGURE 23. **End to end delay vs simulation time - Services D comparison low congestion.**

TABLE 5. **Best effort vs low congestion TSN.**

	A service	B service	C service	D service
BE avg[6]	0,003419	0,000948	0,001892	0,000728
BE std [7]	0,000088	0,000327	0,000045	0,000467
TSN avg	0,003419	0,000949	0,0018	0,000728
TSN std	0,000088	0,000327	0,000033	0,000467
Analysis avg	0,00%	-0,11%	-4,92%	0,00%
Analysis std	0,00%	-0,92%	-26,67%	0,00%

[6] Average.

[7] Standard deviation.

TABLE 6. **Best effort vs low congestion EDGE 1.**

	A service	B service	C service	D service
BE avg	0,003419	0,000948	0,001892	0,000728
BE std	0,000088	0,000327	0,000045	0,000467
EDGE 1 avg	0,003419	0,000948	0,0018	0,000468
EDGE 1 std	0,000088	0,000327	0,000036	0,000222
Analysis avg	0,00%	0,00%	-4,86%	-35,71%
Analysis std	0,00%	0,00%	-20,00%	-52,46%

2) COMPARISON OF SCENARIOS

To carry out the qualitative analysis, the Best Effort scenario will be taken as a basis and comparisons will be made with the following scenarios to establish the relevant changes that occurred.

a: BEST EFFORT VS TSN

For this first scenario (Table 5), it is observed that there is a small improvement for service C, since there is a reduction in the average latency time (4.92%) and a reduction in the standard deviation (26%). This is mainly due to the application of queuing policies. What is interesting in this case is that there is no improvement for service D, which has the highest priority, which means that the saturation in the topology is very low and there is no relevant impact between the traffic generated by the different services.

b: BEST EFFORT VS EDGE 1

In this comparison (Table 6), we observe an improvement in service D, this is mainly due to the reduction of hops within the topology to reach the destination, whereby a reduction of 35.71% in latency time and 52.46% in standard deviation is observed. . By observing the other traffic, the behavior found with Best Effort is maintained, which means that there is still

TABLE 7. **Best effort vs low congestion EDGE 2.**

	A service	B service	C service	D service
BE avg	0,003419	0,000948	0,001892	0,000728
BE std	0,000088	0,000327	0,000045	0,000467
EDGE 2 avg	0,003419	0,000948	0,0018	0,000255
EDGE 2 std	0,000088	0,000327	0,000036	0,000006
Analysis avg	0,00%	0,00%	-4,86%	-64,97%
Analysis std	0,00%	0,00%	-20,00%	-98,72%

TABLE 8. **TSN vs low congestion EDGE 1.**

	A service	B service	C service	D service
TSN avg	0,003419	0,000949	0,001799	0,000728
TSN std	0,000088	0,00033	0,000033	0,000467
EDGE 1 avg	0,003419	0,000948	0,0018	0,000468
EDGE 1 std	0,000088	0,000327	0,000036	0,000222
Analysis avg	0,00%	-0,11%	0,06%	-35,71%
Analysis std	0,00%	-0,91%	9,09%	-52,46%

TABLE 9. **TSN vs low congestion EDGE 2.**

	A service	B service	C service	D service
TSN avg	0,003419	0,000949	0,001799	0,000728
TSN std	0,000088	0,00033	0,000033	0,000467
EDGE 2 avg	0,003419	0,000948	0,0018	0,000255
EDGE 2 std	0,000088	0,000327	0,000036	0,000006
Analysis avg	0,00%	-0,11%	0,06%	-64,97%
Analysis std	0,00%	-0,91%	9,09%	-98,72%

no relevant impact between the different traffic that transits the topology.

c: BEST EFFORT VS EDGE 2

Once again, an improvement is observed for service D (Table 7), due to the reduction of the hops that must be executed within the topology to arrive at the destination, wherey, in this case, we have a reduction of 64.97% and a latency reduction of 98.72 %. For this case, the variation of the latency of service D is given exclusively by the volume of traffic of this service, since the attention nodes are in the edge switches of the topology. The same behavior is maintained for the other services.

d: TSN VS EDGE 1

In this comparison (Table 8), it is observed that the trend evidenced in the other simulations is maintained; that is, a reduction in the latency time of service D, and a similar behavior for the others. Interestingly, an increasing time is observed for Service C; however, since it is such a low variation (0.06%), it is considered an atypical case of the simulation, which does not affect the trend that has been observed.

e: TSN VS EDGE 2

For the comparison of TSN with EDGE 2 (Table 9), the expected improvement is observed in service D and for the other services, whereby it can be concluded that the traffic of service D under this level of saturation is not generating relevant affectation for the other services.

TABLE 10. **Low congestion EDGE 1 vs low congestion EDGE 2.**

	A service	B service	C service	D service
EDGE 1 avg	0,003419	0,000948	0,0018	0,000468
EDGE 1 std	0,000088	0,000327	0,000036	0,000222
EDGE 2 avg	0,003419	0,000948	0,0018	0,000255
EDGE 2 std	0,000088	0,000327	0,000036	0,000006
Analysis avg	0,00%	0,00%	0,00%	-45,51%
Analysis std	0,00%	0,00%	0,00%	-97,30%

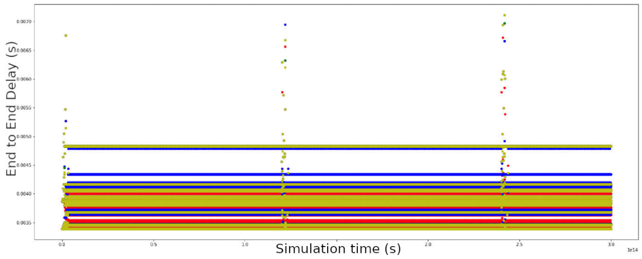

FIGURE 24. **End to end delay vs simulation time - a services comparison - mid congestion.**

f: EDGE 1 VS EDGE 2

In Table 10, an consistent improvement with what is expected is observed for the services that have higher priority, since by including TSN technologies (traffic queuing through credits (CBS) and the ability to prioritize packets at the point of interrupting transmission of low priority traffic), lower latency times and variability for critical services are experimented. Considering this level of saturation, no relevant negative effects are observed for low priority traffic.

B. MID CONGESTION

1) COMPARATIVE ANALYSIS OF SERVICES

Taking into account all the scenarios proposed, a comparative analysis is carried out for each service:

- Service A: Figure 24 shows the simulation results considering the service A for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). Under this scenario, it is observed that the optimization processes begin to affect the latency behavior for this service. In Figure 24, it can be seen that a significant portion of the traffic of this service in the EDGE 2 scenario has a higher latency time, due to the fact that the traffic of this application, being of lower priority, is affected by the traffic of the other applications.
- Service B: Figure 25 shows the simulation results considering the service B for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). Similar to the behavior of service A, service B presents higher latencies in the scenarios with TSN and EDGE optimization processes, although the variation is smaller compared to the effect on service A, this is due again to the priorities,

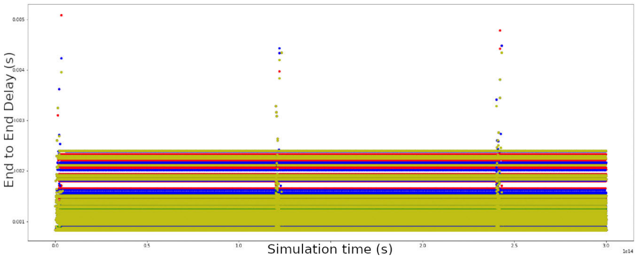

FIGURE 25. End to end delay vs simulation time - B services comparison - mid congestion.

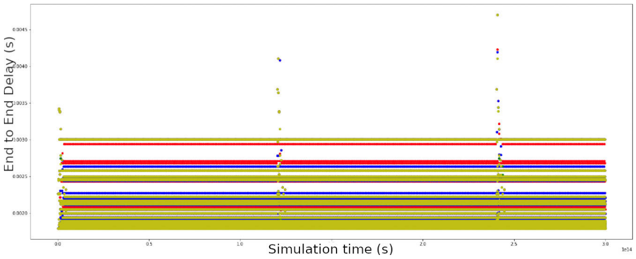

FIGURE 26. End to end delay vs simulation time - C services comparison - mid congestion.

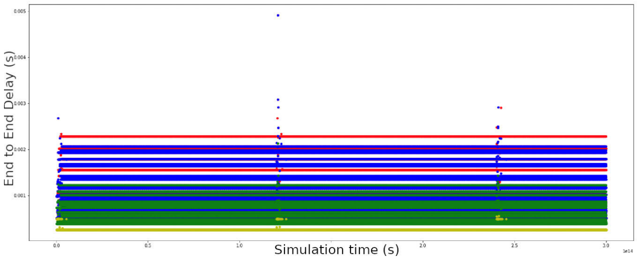

FIGURE 27. End to end delay vs simulation time - D services comparison - mid congestion.

but the effect is lower since there are only two applications with higher priority than application B.

- Service C: Figure 26 shows the simulation results considering the service C for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). In service C, like the previous ones, it presents a variation in its behavior, but in this case it tends to be positive, although there are some peaks in its value. In general, an improvement in latency is observed for the scenarios where techniques of optimization of TSN and EDGE are implemented.
- Service D: Figure 27 shows the simulation results considering the service D for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). As expected, this service presents substantial improvements in latency as the optimization processes are applied, reaching stable latency levels in the EDGE 2 scenario. It is important to emphasize that the expected behavior is maintained even with a medium level of congestion.

TABLE 11. Best effort vs mid congestion TSN.

	A service	B service	C service	D service
BE avg	0,003553	0,001118	0,001917	0,000898
BE std	0,000292	0,000404	0,000252	0,000539
TSN avg	0,00362	0,001054	0,001892	0,000827
TSN std	0,000358	0,000322	0,000218	0,000467
Analysis avg	1,89%	-5,72%	-1,30%	-7,91%
Analysis std	22,60%	-20,30%	-13,49%	-13,36%

TABLE 12. Best effort vs mid congestion EDGE 1.

	A service	B service	C service	D service
BE avg	0,003553	0,001118	0,001917	0,000898
BE std	0,000292	0,000404	0,000252	0,000539
EDGE 1 avg	0,003588	0,001035	0,00188	0,000492
EDGE 1 std	0,000323	0,000305	0,000214	0,000206
Analysis avg	0,99%	-7,42%	-1,93%	-45,21%
Analysis std	10,62%	-24,50%	-15,08%	-61,78%

2) COMPARISON OF SCENARIOS

To carry out the qualitative analysis, the Best Effort scenario will be taken as a basis and comparisons will be made with the following scenarios to establish the relevant changes that occurred.

a: BEST EFFORT VS TSN

By observing the results of the simulation in medium congestion (Table 11), we find that interesting behaviors are already beginning to be observed in the services. For service D, a latency reduction of 7.91% is observed using TSN against Best Effort, which implies that, under this level of saturation, the traffic of the different services is generating an impact on the total transmission times, mainly in the queuing times. However, service A presented a worse behavior using TSN due to the fact that, being the traffic of lower priority, the queuing times are longer, since the queuing attention algorithms are focused on the highest priority queues. The behavior of traffic B is interesting, which has a greater improvement (5.72%) compared to service C (1.30%), despite being a lower priority traffic. This is due to the fact that the traffic of service A, being the one with the lowest priority and the largest packets, reduces its impact on the other services and, therefore, a reduction in the average latency times can be observed.

b: BEST EFFORT VS EDGE 1

In this case (Table 12), the behavior observed in the low congestion scenario is maintained. There is a substantial reduction of the service D (45.21%), due to the reduced amount of hops offered to the topology. At the same time, a greater improvement is observed for services B and C, 7.42% and 1.93% respectively, since by handling priority queues, these traffics can be primarily attended respect to the traffic of service A. For this case, it is observed that the negative impact on service A is reduced (0.99%) compared to the case of Best Effort vs TSN (1.89%), since when downloading the traffic of service D in a closer node, it is reduced the impact

TABLE 13. Best effort vs mid congestion EDGE 2.

	A service	B service	C service	D service
BE avg	0,003553	0,001118	0,001917	0,000898
BE std	0,000292	0,000404	0,000252	0,000539
EDGE 2 avg	0,003582	0,001041	0,001879	0,000255
EDGE 2 std	0,000325	0,000319	0,000214	0,000006
Analysis avg	0,82%	-6,89%	-1,98%	-71,60%
Analysis std	11,30%	-21,04%	-15,08%	-98,89%

TABLE 14. TSN vs mid congestion EDGE 1.

	A service	B service	C service	D service
TSN avg	0,00362	0,001054	0,001892	0,000827
TSN std	0,000358	0,000322	0,000218	0,000467
EDGE 1 avg	0,003588	0,001035	0,00188	0,000492
EDGE 1 std	0,000323	0,000305	0,000214	0,000206
Analysis avg	-0,88%	-1,80%	-0,63%	-40,51%
Analysis std	-9,78%	-5,28%	-1,83%	-55,89%

TABLE 15. TSN vs mid congestion EDGE 2.

	A service	B service	C service	D service
TSN avg	0,00362	0,001054	0,001892	0,000827
TSN std	0,000358	0,000322	0,000218	0,000467
EDGE 2 avg	0,003582	0,001041	0,001879	0,000255
EDGE 2 std	0,000325	0,000319	0,000214	0,000006
Analysis avg	-1,05%	-1,23%	-0,69%	-69,17%
Analysis std	-9,22%	-0,93%	-1,83%	-98,72%

of this traffic on the bottleneck of the proposed topology (SW8), whereby service A has an improvement in its latency averages.

c: BEST EFFORT VS EDGE 2

In Table 13, the comparison of scenarios maintains the same trend as the previous one, with very similar values. The only relevant difference is the reduction in the average latency times for service D. This is due to the fact that, in this case, by bringing the attention nodes of service D even closer, only the times associated with this service are reduced, but the impact on the other services is minimal since, in this case, the topology does not present a bottleneck as clear as it was in the base case (SW8).

d: TSN VS EDGE 1

This case (Table 14) presents favorable results for all services, including service A, with a reduction of 0.88% in the average latency. This is due to the fact that by moving attention from service D to Switches Sw5 and Sw7, it is deleted the impact of service D on the bottlenecks that the topology had, which were the connections from Switches Sw5 and Sw7 to Switch Sw8. Additionally, the behavior of service C is interesting, since although there is an improvement, it is not significant (0.63%). In this sense, it can be established that under these levels of congestion, service D was not generating a relevant impact on the others services.

e: TSN VS EDGE 2

For this comparison (Table 15), similar values to what had been observed were maintained. There is a slight additional improvement for service C (0.69%) compared to the EDGE 1 model (0.63%); however, this means a lower reduction for service B, which changed from 1.80% in EDGE 1 to 1.23% in EDGE2. Finally, for service A, there is an additional improvement since it changed from 0.88% to 1.05% in average latency reductions. All these improvements are caused by the removal of the impact of service D traffic on the topology, whereby the service C becomes the one with the highest priority after assuming edge switches, reducing the impact on the bottlenecks presented in the topology.

TABLE 16. EDGE 1 vs mid congestion EDGE 2.

	A service	B service	C service	D service
EDGE 1 avg	0,003588	0,001035	0,00188	0,000492
EDGE 1 std	0,000323	0,000305	0,000214	0,000206
EDGE 2 avg	0,003582	0,001041	0,001879	0,000255
EDGE 2 std	0,000325	0,000319	0,000214	0,000006
Analysis avg	-0,17%	0,58%	-0,05%	-48,17%
Analysis std	0,62%	4,59%	0,00%	-97,09%

f: EDGE 1 VS EDGE 2

In Table 16, for the EDGE1 and EDGE2 models, excluding service D, there are no significant changes, since from the EDGE 1 model, the impact of service D traffic on topology bottlenecks was eliminated. Thus, in this scenario, only a substantial improvement of service D is observed, by reducing the number of hops to reach the attention node. The trend of optimizing latency times for priority services is maintained, considering that network saturation has increased. Additionally, it is observed that the implementation of EDGE schemes reduces the impact on low priority services by reducing the saturation on the critical links of the topology.

C. HIGH CONGESTION

1) COMPARATIVE ANALYSIS OF SERVICES

Taking into account all the scenarios proposed, a comparative analysis is carried out for each of the services:

- Service A: Figure 28 shows the simulation results considering the service A for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). Under these conditions, it is observed that the EDGE optimization schemes contribute to the behavior of this service indirectly, since by simplifying the attention of priority services, the pressure on the network topology is unloaded, whereby services with lower priority as in this case, have more resources available and can reduce their latency.
- Service B: Figure 29 shows the simulation results considering the service B for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). In this case, the results are clearer to visualize the effect of the TSN and EDGE optimization processes, since it is observed that the distribution for the EDGE 2 scenario have lower latency values and less variation throughout the simulation.

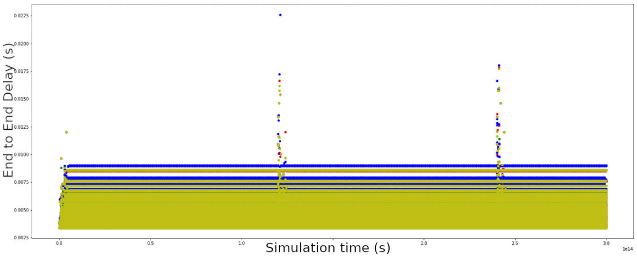

FIGURE 28. End to end delay vs simulation time - A services comparison - high congestion.

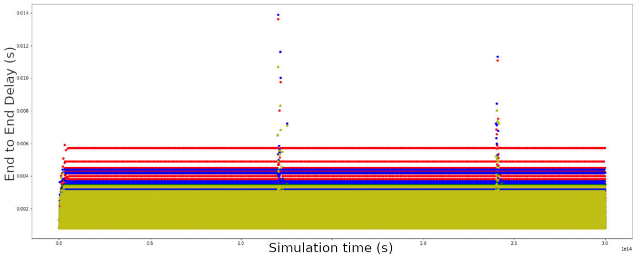

FIGURE 29. End to end delay vs simulation time - B services comparison - high congestion.

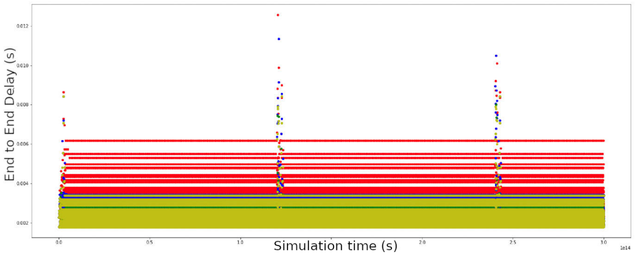

FIGURE 30. End to end delay vs simulation time - C services comparison - high congestion.

- Service C: Figure 30 shows the simulation results considering the service C for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). In this particular case, there is a clear advantage over the behavior between Best Effort and EDGE 2, with a reduction in latency times close to 50%. This makes sense, since by preventing the priority services from travelling through the entire topology (this, being attended as close as possible), the rest of the resources are made available for the other services and, in this case, service C is the one that still maintains priority levels.
- Service D: Figure 31 shows the simulation results considering the service D for Best Effort, TSN, Edge optimization 1 and Edge optimization 2 technologies (red, blue, green and yellow, respectively). In this service, the optimization levels that are generated with each of the implemented stages are clearly observed, where the Best Effort model delivers the worst results. EDGE 2 model maintains low latency levels and very little variability.

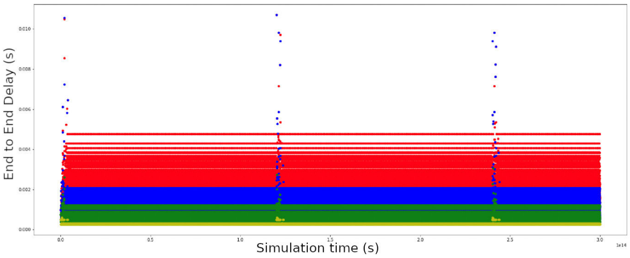

FIGURE 31. End to end delay vs simulation time - D services comparison - high congestion.

TABLE 17. Best effort vs high congestion TSN.

	A service	B service	C service	D service
BE avg	0,004181	0,001523	0,002464	0,001199
BE std	0,00093	0,000713	0,000753	0,000779
TSN avg	0,004349	0,001378	0,002176	0,000853
TSN std	0,001077	0,000571	0,000427	0,000391
Analysis avg	4,02%	-9,52%	-11,69%	-28,86%
Analysis std	15,81%	-19,92%	-43,29%	-49,81%

TABLE 18. Best effort vs high congestion EDGE 1.

	A service	B service	C service	D service
BE avg	0,004181	0,001523	0,002464	0,001199
BE std	0,00093	0,000713	0,000753	0,000779
EDGE 1 avg	0,004233	0,001323	0,002127	0,000477
EDGE 1 std	0,000996	0,00052	0,000391	0,000202
Analysis avg	1,24%	-13,13%	-13,68%	-60,22%
Analysis std	7,10%	-27,07%	-48,07%	-74,07%

2) COMPARISON OF SCENARIOS

To carry out the qualitative analysis, the Best Effort scenario will be taken as a basis and comparisons will be made with the following scenarios to establish the relevant changes that occurred.

a: BEST EFFORT VS TSN

In Table 17, in the high congestion scenario, we find that TSN begins to take a relevant advantage over Best Effort, since it presents a reduction of 25.86% for service D, 11.69% for service C and 9.52% for service B, all this maintaining the same deployment topology, but optimizing the resource allocation, queuing and packet breaking algorithms. These results have a negative impact on service A, whose average latency increased by 4.02% and its standard deviation grew by 15.81%. This result shows the potential of TSN for high traffic cases and how a correctly designed deployment allows us to ensure the service parameters for multiple services with a reduced impact on less critical services.

b: BEST EFFORT VS EDGE 1

For the scenario using the EDGE 1 topology (Table 18), an additional improvement is observed on services C and B of around 13% compared to Best Effort, and a reduction in the negative impact on service A from 4.02%, in the comparison with TSN, to 1.24% using TSN and EDGE technologies.

TABLE 19. Best effort vs high congestion EDGE 2.

	A service	B service	C service	D service
BE avg	0,004181	0,001523	0,002464	0,001199
BE std	0,00093	0,000713	0,000753	0,000779
EDGE 2 avg	0,004221	0,001313	0,002125	0,000256
EDGE 2 std	0,000994	0,000525	0,000391	0,000008
Analysis avg	0,96%	-13,79%	-13,76%	-78,65%
Analysis std	6,88%	-26,37%	-48,07%	-98,97%

TABLE 20. TSN vs high congestion EDGE 1.

	A service	B service	C service	D service
TSN avg	0,004349	0,001378	0,002176	0,000853
TSN std	0,001077	0,000571	0,000427	0,000391
EDGE 1 avg	0,004233	0,001323	0,002127	0,000477
EDGE 1 std	0,000996	0,00052	0,000391	0,000202
Analysis avg	-2,67%	-3,99%	-2,25%	-44,08%
Analysis std	-7,52%	-8,93%	-8,43%	-48,34%

An interesting value to take into account in this comparison is the reduction of almost 50% of the standard deviation of the latency of service C, which presents much more stable transmission times even with high levels of congestion. Finally, for service D, there is a reduction of 60% in the average latency, and 74% in the standard deviation, which makes it possible to ensure fairly strict network parameters for this service.

c: BEST EFFORT VS EDGE 2

For this comparison scenario (Table 19), the values for services C, B and A, which were obtained in the EDGE 1 topology, are maintained with some minor variations since, as has been observed in the other saturation scenarios, topology changes from EDGE 1 to EDGE 2 do not generate substantial changes to services (except D service). In the case of service D, the results obtained are quite interesting in terms of standard deviation of latency, since a reduction of 98.97% is achieved, which implies that, regardless of having a high congestion scenario, the latency times remain stable, which allows reliability on the required network parameters under different levels of congestion.

d: TSN VS EDGE 1

When comparing TSN with EDGE 1 (Table 20), it is observed that there is a reduction in all the services deployed, which implies that, under this level of topology saturation, there is a tangible effect of the traffic of service D on the other services deployed. For this reason, a displacement of the attention nodes is experimented, freeing the critical points SW5 and SW7 towards SW8, providing a better behavior of services C, B and A.

e: TSN VS EDGE 2

For this scenario (Table 21), there are no relevant changes compared to the previous comparison, except for the reduction of the metrics of service D, in particular, the standard deviation of latency, caused by the attention nodes of service D are being located in the edge Switches SW1 - SW4.

TABLE 21. TSN vs high congestion EDGE 2.

	A service	B service	C service	D service
TSN avg	0,004349	0,001378	0,002176	0,000853
TSN std	0,001077	0,000571	0,000427	0,000391
EDGE 2 avg	0,004221	0,001313	0,002125	0,000256
EDGE 2 std	0,000994	0,000525	0,000391	0,000008
Analysis avg	-2,94%	-4,72%	-2,34%	-69,99%
Analysis std	-7,71%	-8,06%	-8,43%	-97,95%

TABLE 22. EDGE 1 vs high congestion EDGE 2.

	A service	B service	C service	D service
EDGE 1 avg	0,004233	0,001323	0,002127	0,000477
EDGE 1 std	0,000996	0,00052	0,000391	0,000202
EDGE 2 avg	0,004221	0,001313	0,002125	0,000256
EDGE 2 std	0,000994	0,000525	0,000391	0,000008
Analysis avg	-0,28%	-0,76%	-0,09%	-46,33%
Analysis std	-0,20%	0,96%	0,00%	-96,04%

f: EDGE 1 VS EDGE 2

In Table 22, the results present a similar behavior. The use of the optimization models on the topology improves the latency metrics of the critical services. However, the adverse effect on secondary services is mitigated by implementing EDGE computing schemes, since it is possible to free network resources that can be used for low priority traffic, mitigating and even improving the performance against Best Effort schemes.

D. QUEUEING, LATENCY AND JITTER RESULTS

For this subsection, it was performed two scenarios: Best Effort and TSN (including traffic programming based on credits). In addition, as we mentioned in the implementation section (IV), the results were obtained considering three congestion levels. Finally, for obtaining these results, it was used Omnet 6.0 with the package INET 4.4. More details are described as follows:

- Best Effort: basically all traffic has the same priority and use the same queue.
- General scheme for TSN:
  - Network synchronization: 802.1As
  - Flows filtering: 802.Qci
  - Time Aware Shaper: 802.1 Qbv
  - Gates configuration: 802.1Qcc
  - Traffic programming based on credits: 802.1Qav
  - Frame Preemption (802.1Qbu)

Based on the need to evaluate additional parameters on the behavior of the network in TSN technologies, an extension of the simulation processes was proposed in which it is possible to demonstrate additional parameters on the traffic using different available traffic schedulers.

Within this extension of the simulation, the comparative analysis of two technologies is established:

- Best Effort: In this case, it is assumed that all traffic is handled with the same priority.
- CBS (Credit Bases Shapper): for the second scenario, the use of a credit-based programmer is considered.

TABLE 23. Applications for the extended simulation.

ID	Description	Packet size	Priority (PCP)
1	Priority traffic for IoT devices states	500 Bytes	6
2	Priority traffic for IoT devices states	200 Bytes	5
3	Priority traffic for IoT devices states	1400 Bytes	4
4	Priority traffic for IoT devices states	1000 Bytes	0

According to the applications to be modeled, we present them in Table 23.

Within each scenario, three saturation levels are defined to evaluate the behavior of the network indicators.

- Low: an approximate utilization of 30%-32% is established.
- Medium: an approximate utilization of 48% - 52% is established.
- High: an approximate utilization of 80% is established.

The volume of data generated by each of the applications is distributed uniformly, to ensure that the behavior variation is given by the traffic management that is generated within the aggregation switch.

The network parameters to be taken into account in the simulations are the following:

- Latency: the latency behavior of each of the types of applications mentioned will be analyzed.
- Jitter: the behavior of latency variation in packets will be analyzed in order to assess the impact of traffic or services sensitive to this type of variation, such as the transmission of multimedia services.
- Queuing: to visualize more clearly the behavior of the queues for each of the traffics and to evaluate the impact of the different traffic scheduling schemes that are being used.

Finally, the topology to be used (see Figure 51) was simplified to focus the results, where a central Switch with TSN capacity and four traffic sources is proposed to simulate the different applications and a single receiver, in order to force the queuing process in the transmission link communications:

Regarding the saturation levels and the different services that travel in the topology, the following values were proposed (see Tables 24, 25 and 26).

1) BEST EFFORT SCENARIO

a: LOW SATURATION

Queue behavior (see Figure 52 and Table 27):

A stable behavior of the queue is observed after a high growth at the beginning of the simulation with some peaks during the execution.

Reviewing to these results, a peak of 29 packets is observed at the beginning of the simulation with a stable queuing around 3 packets.

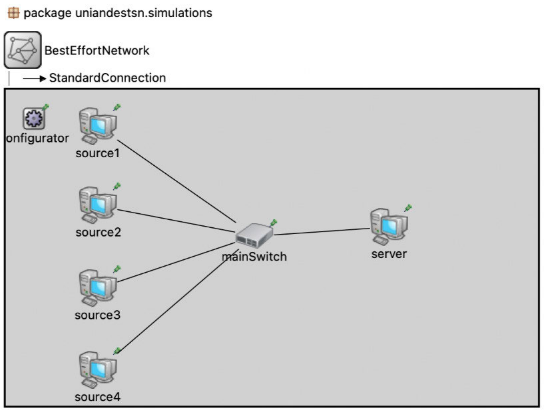

FIGURE 32. Simulation extended topology.

TABLE 24. Configuration for low saturation applications.

Low Saturation: level	Critic	Random	Video	No priority
Packet size (Byte)	500	200	1400	1000
Packet size (Bit)	4000	1600	11200	8000
Frequency (Packet/sec)	2000	5000	714,28	1000
Interval generation (sec)	0,0005	0,0002	0,0014	0,001
Real (s) interval	0,0005	0,0002	0,0014	0,001
Real bandwidth (b/s)	8000000	8000000	8000000	8000000
Bandwidth (b/s) (b/s)	8000000	8000000	8000000	8000000

TABLE 25. Configuration for mid saturation applications.

Mid saturation: level	Critic	Random	Video	No priority
Packet size (Byte)	500	200	1400	1000
Packet size (Bit)	4000	1600	11200	8000
Frequency (Packet/sec)	3000	7500	1071,42	1500
Interval generation (sec)	0,00033333	0,00013333	0,00093333	0,00066667
Real (s) interval	0,0003	0,0001	0,0009	0,0006
Real bandwidth (b/s)	13333333,3	16000000	12444444,4	13333333,3
Bandwidth (b/s) (b/s)	12000000	12000000	12000000	12000000

Latency (see Figure 34):

Figure 34 shows latency results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). The traffic latency for the different applications is similar, with small variations since the same attention queue is handled and channel saturation is low.

Jitter (see Figure 35):

Figure 35 shows jitter results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). The Jitter behavior remains stable during

TABLE 26. Configuration for high saturation applications.

High saturation: level	Critic	Random	Video	No priority
Packet size (Byte)	500	200	1400	1000
Packet size (Bit)	4000	1600	11200	8000
Frequency (Packet/sec)	5000	12500	1785,71	2500
Interval generation (sec)	0,0002	0,00008	0,00056	0,0004
Real (s) interval	0,0002	0,00008	0,0005	0,0004
Real bandwidth (b/s)	20000000	20000000	22400000	20000000
Bandwidth (b/s) (b/s)	20000000	20000000	20000000	20000000

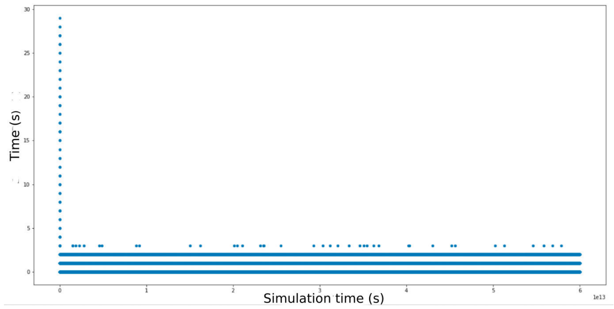

FIGURE 33. Queue behavior for best effort with low congestion.

TABLE 27. Queue behavior for best effort with low congestion.

	simtimeRaw	value
count	6.651300e+05	665130
mean	2.999544e+ 13	0.249163
std	1.732404e+ 13	0.509412
min	1.010760e+09	0
25%	1.498714e+13	0
50%	2.999516e+13	0
75%	4.500027e+13	0
max	5.999986e+13	29

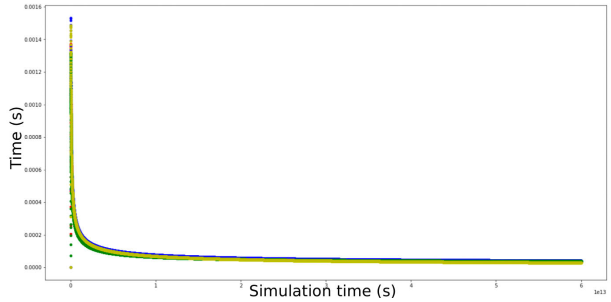

FIGURE 34. Latency behavior for best effort with low congestion.

the simulation for the four applications, because the use of the channel is low and there is no high queuing processes.

b: MID SATURATION

Queue behavior (see Figure 36 and Table 28):

In this scenario, a growth in the packet queuing is obtained, reaching values of approximately 5 packets.

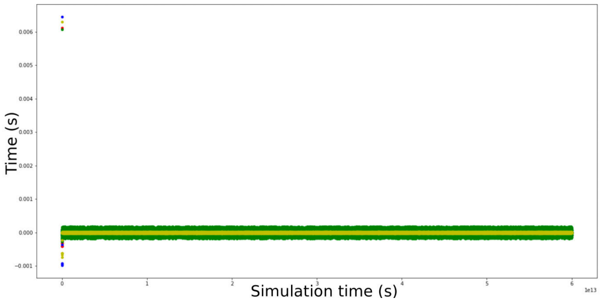

FIGURE 35. Jitter for best effort with low congestion.

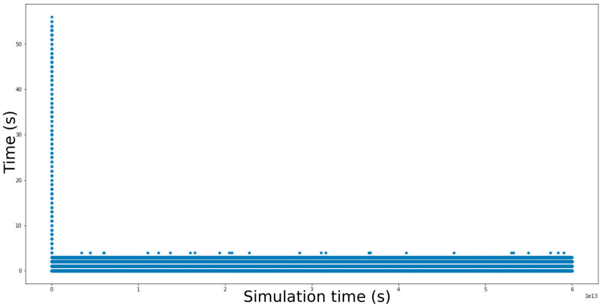

FIGURE 36. Queue behavior for best effort with mid congestion.

TABLE 28. Queue behavior for best effort with mid congestion.

	simtimeRaw	value
count	1.442135e+06	1.442135e+06
mean	3.000072e+13	4.910158e-01
std	1.732917e+13	8.285799e-01
min	1.010760e+09	0
25,00%	1.498816e+13	0
50,00%	2.999899e+13	0
75,00%	4.501603e+13	1
max	5.999995e+13	5.60e+01

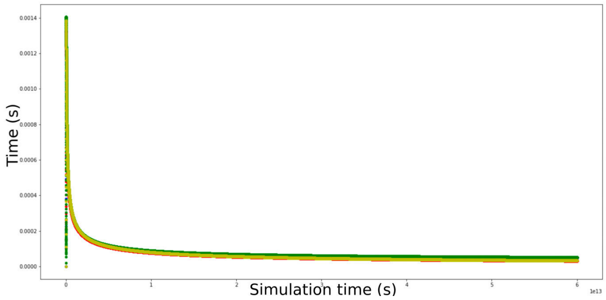

FIGURE 37. Latency for best effort with mid congestion.

Analyzing these results, a peak is observed at the beginning of the simulation and a stabilization as progress is made in the data, maintaining a value close to 5, although there is an edge effect at the beginning of the simulation, which affects part of the statistical calculated results.

Latency (see Figure 37):

Figure 37 shows latency results considering critical, random, multimedia and no prioritized traffic (red, green, blue

TABLE 29. Latency for best effort with mid congestion and critical traffic.

	simtimeRaw	value
count	1.999920e+05	199992
mean	3.000075e+13	0.000062
std	1.731986e+13	0.000063
min	6.127400e+09	0
25%	1.500142e+13	0.000037
50%	3.000075e+13	0.000045
75%	4.500008e+13	0.000063
max	5.999940e+13	0.001324

TABLE 30. Latency for best effort with mid congestion and random traffic.

	simtimeRaw	value
count	6.001400e+05	600140
mean	3.000814e+13	0.000077
std	1.732738e+ 13	0.000063
min	6.082120e+09	0
25%	1.499319e+13	0.000055
50%	3.001723e+13	0.000061
75%	4.502273e+13	0.000076
max	5.999995e+13	0.001405

TABLE 31. Latency for best effort with mid congestion and multimedia traffic.

	simtimeRaw	value
count	6.666400e+04	66664
mean	3.000063e+13	0.000076
std	1.731994e+13	0.000063
min	6.460360e+09	0
25%	1.500147e+13	0.000053
50%	3.000063e+13	0.000059
75%	4.499979e+13	0.000075
max	5.999895e+13	0.001376

TABLE 32. Latency for best effort with mid congestion and no prioritized traffic.

	simtimeRaw	value
count	9.999600e+04	99996
mean	3.000070e+13	0.000065
std	1.731990e+ 13	0.000066
min	6.321800e+09	0
25,00%	1.500145e+13	0.000038
50,00%	3.000069e+ 13	0.000047
75,00%	4.499995e+13	0.000066
max	5.999919e+13	0.001384

and yellow, respectively). The latency behavior is maintained compared to the low saturation levels observed; however, a higher latency trend is observed for random traffic packets, probably associated with the fact that there is more queuing and this traffic is more affected by the queuing growth.

According to the Tables 29, 30, 31 and 32, the latency of random traffic is higher than other traffic. This is mainly due to the increase in queuing and the low generation rate compared to other traffic, which makes it more susceptible to edge effects, as those presented at the beginning of the simulation.

Jitter (see Figure 38):

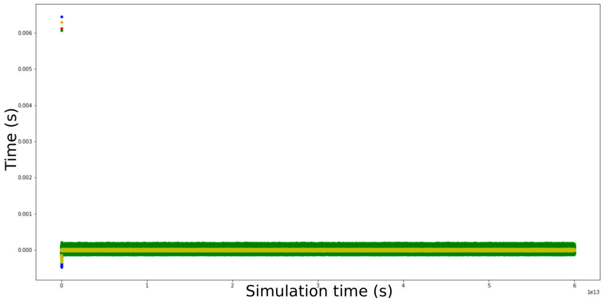

FIGURE 38. Jitter for best effort with mid congestion.

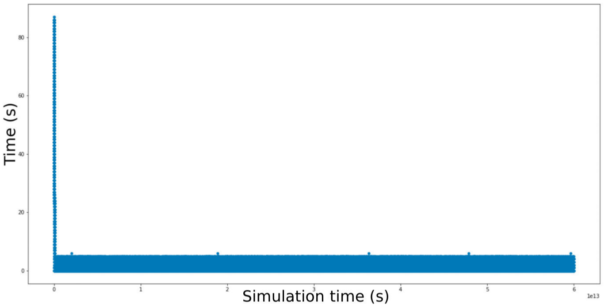

FIGURE 39. Queue behavior for best effort with high congestion.

TABLE 33. Queue behavior for best effort with mid congestion.

	simtimeRaw	value
count	2.440665e+06	2.440665e+06
mean	3.000227e+13	1.24
std	1.732304e+13	1.98
min	1.010760e+09	0
25,00%	1.499529e+13	0
50,00%	3.001364e+13	1
75,00%	4.500345e+13	2
max	5.999995e+13	87

Figure 38 shows jitter results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). An increase in jitter is observed compared to low congestion levels, although the values remain stable during the simulation, due to the higher levels of queuing. In thise sense, A growth in the Jitter of the applications is obtained, expected according to the growth of the channel saturation in all the applications that flow in the topology.

c: HIGH SATURATION

Queue behavior (see Figure 39 and Table 33):

For the high congestion scenario, the edge effect at the beginning of the simulation generates a saturation of the queue and, after that, it stabilizes and peaks close to 10 are presented.

In Table 33 there is a maximum of 87 packages at the beginning of the simulation and, after that, the behavior stabilizes with an average of 1.2 packages with a deviation close to 2 packages.

Latency (see Figure 40):

Figure 40 shows results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). The same behavior is observed as the medium and low levels of queuing, and the effect that it had on random

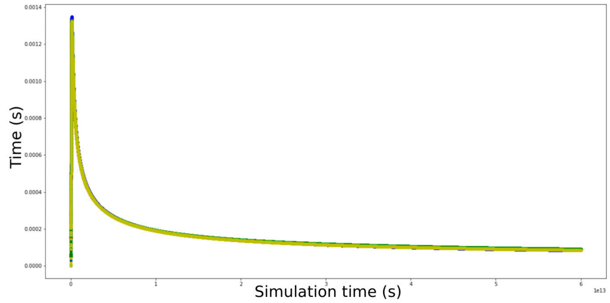

FIGURE 40. Latency for best effort with high congestion.

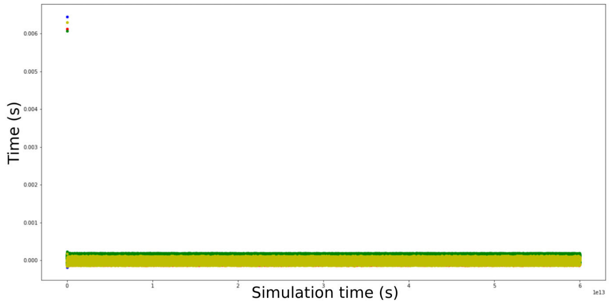

FIGURE 41. Jitter for best effort with high congestion.

traffic is reduced since the packet generation rate is increased, mitigating the effect that was observed in the medium congestion. In this sense, the results maintains the trend observed in the previous scenarios, with a similar latency for the four types of traffic that flow in the topology.

Jitter (see Figure 41):

Figure 41 shows jitter results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). Following the trend in relation with the previous results, a similar behavior to the other congestion levels is observed, but a higher distribution is observed in the jitter values, associated with the queuing process.

2) CBS SCENARIO

For the scenarios under the TSN - CBS model, the idle slopes are established for each of the traffics that flow through the topology in order to establish the growth and decrease of the allocation credit.

a: LOW CONGESTION

For the simulation using CBS, the same topology and types of traffic are used with the following *IdleSlope* parameters for the traffic:

- Priority Traffic: 16 Mbps
- Random Traffic: 12 Mbps
- Multimedia Traffic: 12 Mbps
- Non Relevant Traffic: 12 Mbps

Queue behavior (see Figure 42 and Tables 34, 35, 36 and 37):

In the first scenario, it is observed that there is no queuing in the different queues of the main switch. For this reason, this simulation is executed several times to ensure consistency.

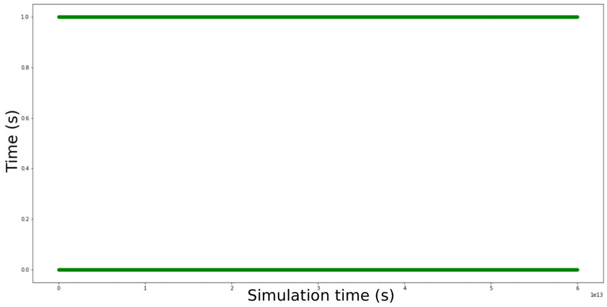

FIGURE 42. Queue behavior for CBS with low congestion.

TABLE 34. Queue behavior for CBS with low congestion and critical traffic.

	simtimeRaw	value
count	1.398160e+05	139816
mean	2.999732e+13	0.141729
std	1.731880e+13	0.348773
min	3.083100e+08	0
25,00%	1.500892e+13	0
50,00%	2.999929e+13	0
75,00%	4.498667e+13	0
max	5.999955e+13	1

TABLE 35. Queue behavior for CBS with low congestion and random traffic.

	simtimeRaw	value
count	2.523290e+05	252329
mean	2.999954e+13	0.207269
std	1.731622e+13	0.405351
min	2.843100e+08	0
25,00%	1.501131e+13	0
50,00%	2.998970e+13	0
75,00%	4.499888e+13	0
max	5.999975e+13	1

TABLE 36. Queue behavior for CBS with low congestion and multimedia traffic.

	simtimeRaw	value
count	5.385700e+04	53857
mean	3.000769e+13	0.204226
std	1.732503e+13	0.403139
min	3.803100e+08	0
25,00%	1.500112e+13	0
50,00%	3.002172e+13	0
75,00%	4.502972e+13	0
max	5.999992e+13	1

The statistical analysis corroborates the figure where queuing is observed in the queues for each of the defined traffics.

Latency (see Figure 43):

Figure 43 shows latency results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). In the case of latency, an initial peak associated with the activation of the simulation is observed, and after that, stability is observed with the latency times directly related to the size of the processed traffic packet. As expected, the variation is very small in latency

TABLE 37. Queue behavior for CBS with low congestion and no prioritized traffic.

	simtimeRaw	value
count	1.199990e+05	119999
mean	2.999984e+13	0.499996
std	1.732044e+ 13	0.500002
min	3.599300e+08	0
25,00%	1.500009e+13	0
50,00%	3.000008e+13	0
75,00%	4.499959e+13	1
max	5.999909e+ 13	1

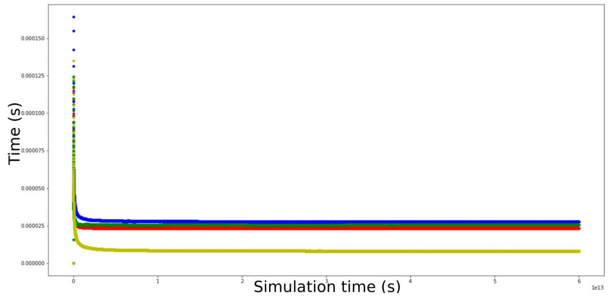

FIGURE 43. Latency for CBS with low congestion.

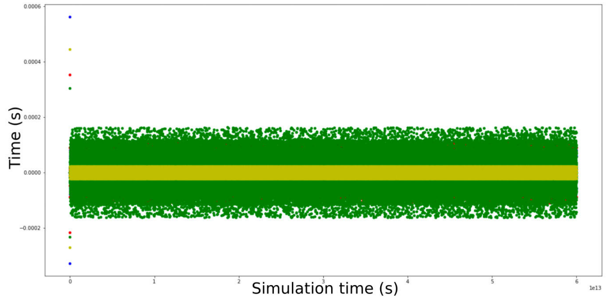

FIGURE 44. Jitter for CBS with low congestion.

TABLE 38. Jitter for CBS with mid congestion and critical traffic.

	simtimeRaw	value
count	1.199990e+05	1.199990e+05
mean	2.999960e+13	7.421729e-10
std	1.732044e+13	3.414415e-05
min	3.526800e+08	-2.170000e-04
25,00%	1.499984e+13	0
50,00%	2.999959e+13	0
75,00%	4.499934e+13	0
max	5.999909e+13	3.526800e-04

times compared to the low levels of queuing presented by the model.

Jitter (see Figure 44 and Tables 38, 39, 40 and 41):

Figure 44 shows jitter results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). In the case of the latency variation, a uniform behavior is observed, which was expected due to the low level of queuing that occurs in the low congestion scenario.

TABLE 39. Jitter for CBS with mid congestion and random traffic.

	simtimeRaw	value
count	2.000280e+05	2.000280e+05
mean	2.998612e+13	2.052713e-10
std	1.731740e+13	3.795253e-05
min	3.046800e+08	-2.320373e-04
25,00%	1.499078e+13	0
50,00%	2.997905e+13	0
75,00%	4.498169e+13	0
max	5.999954e+13	3.046800e-04

TABLE 40. Jitter for CBS with mid congestion and multimedia traffic.

	simtimeRaw	value
count	4.285600e+04	4.285600e+04
mean	2.999875e+13	7.301195e-09
std	1.732025e+13	4.357074e-05
min	5.615800e+08	-3.285200e-04
25,00%	1.499948e+13	0
50,00%	2.999873e+13	0
75,00%	4.499798e+13	0
max	5.999731e+13	5.615800e-04

TABLE 41. Jitter for CBS with mid congestion and no prioritized traffic.

	simtimeRaw	value
count	5.999900e+04	5.999900e+04
mean	2 .999918e+13	2.90571Se-09
std	1.732036e+13	1.157773e-05
min	4.443000e+08	-2.699600e-04
25,00%	1.499968e+13	0
50,00%	2.999917e+13	0
75,00%	4.499867e+13	0
max	5.999817e+13	4.443000e-04

According to Tables 38, 39,40 and 41, the highest variation between the packets is presented by the random and multimedia traffic packets, the first due to its statistical nature and the second associated with the size of the packets, which causes higher attention times.

b: MID CONGESTION

For the simulation using CBS, the same topology and types of traffic are used with the following IdleSlope parameters for the traffic:

- Priority Traffic: 24 Mbps
- Random Traffic: 24 Mbps
- Multimedia Traffic: 18 Mbps
- Non Relevant Traffic: 18 Mbps

Queue behavior (see Figure 45 and Tables 42, 43, 44 and 45):

Under the medium saturation scenario, queuing levels are observed for random traffic, since it can have peaks due to the uniform statistical distribution it handles. For other traffics, no relevant queuing is observed. However, it is important to clarify that priority traffic affects the service of lower priority traffic.

For random traffic, it is important to take into account that service times can be impacted by the generation of higher

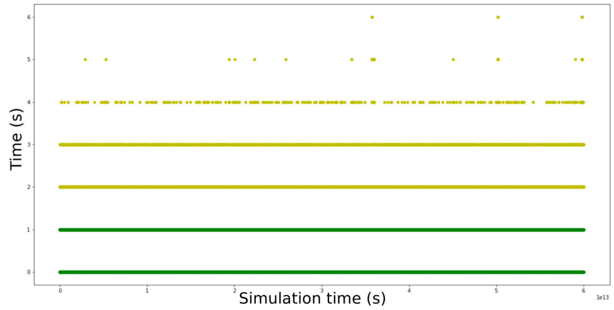

FIGURE 45. Queue behavior for CBS with mid congestion.

TABLE 42. Queue behavior for CBS with mid congestion and critical traffic.

	simtimeRaw	value
count	2.682580e+05	26258
mean	3.000904e+13	0.254.449
std	1.731575e+13	0.435552
min	2.343100e+08	0
25,00%	1.501055e+13	0
50,00%	3.001895e+13	0
75,00%	4.499037e+13	1
max	5.999976e+13	1

TABLE 43. Queue behavior for CBS with mid congestion and random traffic.

	simtimeRaw	value
count	1.064965e+06	1.064965e+06
mean	3.001334e+13	6.209641e-01
std	1.733292e+ 13	6.770461e-01
min	2.893100e+08	0
25,00%	1.498867e+13	0
50,00%	3.002811e+13	1
75,00%	4.503759e+13	1
max	5.999992e+13	6

TABLE 44. Queue behavior for CBS with mid congestion and multimedia traffic.

	simtimeRaw	value
count	1.072270e+05	107227
mean	3.000663e+13	0.378263
std	1.732278e+13	0.484956
min	3.853100e+08	0
25,00%	1.499908e+13	0
50,00%	3.001063e+13	0
75,00%	4.500017e+13	1
max	5.999962e+13	1

priority traffic, in this case critical traffic, which can cause an increase in the number of queued packets.

Latency (see Figure 46):

Figure 46 shows latency results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). For this level of saturation, the benefits of TSN queuing schemes begin to be observed, since although there was an increase in the generated traffic, the prioritization configurations managed to keep the latency time very limited for critical traffic and random traffic, increasing the attention time for lower priority traffic (multimedia and no relevant traffic).

TABLE 45. Queue behavior for CBS with mid congestion and no prioritized traffic.

	simtimeRaw	value
count	1.66371Oe+05	166371
mean	3.001314e+13	0.398934
std	1.733010e+13	0.489681
min	3.817100e+08	0
25,00%	1.500371e+13	0
50,00%	3.001630e+13	0
75,00%	4.502081e+13	1
max	5.999952e+13	1

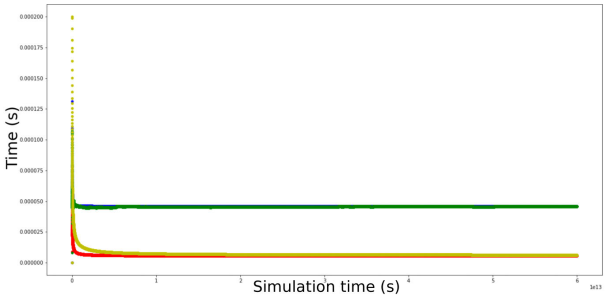

FIGURE 46. Latency for CBS with mid congestion.

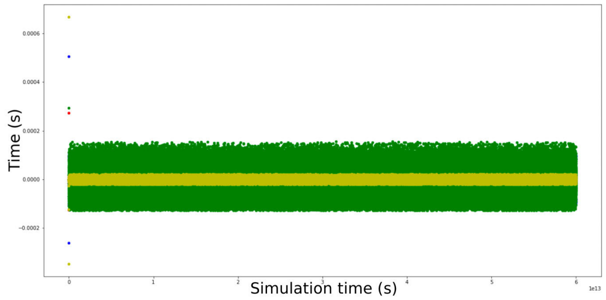

FIGURE 47. Jitter for CBS with mid congestion.

TABLE 46. Jitter for CBS with mid congestion and critical traffic.

	simtimeRaw	value
count	1.999990e+05	1.999990e+05
mean	2.999980e+13	4.453022e-10
std	1.732046e+13	8.383710e-06
min	2.786800e+08	-1.268800e-04
25,00%	1.499995e+13	-1.422717e-06
50,00%	2.999979e+13	0
75,00%	4.499964e+13	1.200000e-07
max	5.999949e+13	2.736800e-04

Jitter (see Figure 47 and Tables 46, 47, 48 and 49):

Figure 34 shows jitter results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). For this level of saturation, a similar behavior is observed compared to medium congestion, which is expected according to the queuing levels obtained. For this reason, the queuing time does not significantly impact the general behavior of traffic.

TABLE 47. Jitter for CBS with mid congestion and random traffic.

	simtimeRaw	value
count	6.001590e+05	6.001590e+05
mean	3.000711e+13	1.337431e-10
std	1.732793e+13	4.490403e-05
min	3.096800e+08	-1.287121e-04
25,00%	1.499172e+13	-2.9205886-05
50,00%	3.001619e+13	0
75,00%	4.502211e+13	2.825526e-05
max	5.999983e+13	2.946800e-04

TABLE 48. Jitter for CBS with mid congestion and multimedia traffic.

	simtimeRaw	value
count	6.666600e+04	6.666600e+04
mean	2.999954e+13	3.495935e-09
std	1.732046e+13	9.116549e-05
min	5.156267e+08	-2.623911e-04
25,00%	1.499994e+13	-9.534592e-05
50,00%	2.999954e+13	8.053500e-09
75,00%	4.499914e+13	9.440056e-05
max	5.999874e+13	5.056267e-04

TABLE 49. Jitter for CBS with mid congestion and no prioritized traffic.

	simtimeRaw	value
count	9.999900e+04	9.999900e+04
mean	2.999960e+13	1.806218e-09
std	1.732042e+13	9.062529e-06
min	6.887467e+08	-3.488889e-04
25,00%	1.499990e+13	-7.215721e-06
50,00%	2.999960e+13	0
75,00%	4.499930e+13	7.170436e-06
max	5.999900e+13	6.687467e-04

According to Tables 46, 47,48 and 49, and respecting the low levels of queuing, the jitter values are relatively low for all traffic, although a growth associated with the greater volume of traffic is observed, and with it, the increasing possibility that the traffic will be received, while the output transmission channel is being used.

c: HIGH CONGESTION

For the simulation using CBS, the same topology and types of traffic are used with the following IdleSlope parameters for the traffic:

- Prioritized traffic: 24 Mbps.
- Random traffic: 24 Mbps.
- Multimedia traffic: 18 Mbps.
- No prioritized: 18 Mbps.

Queue behavior (see Figure 48 and Tables 50, 51, 52 and 53):

The behavior observed in the medium saturation level is maintained. There is queuing of random traffic, although queuing for lower priority traffic is already observed at certain points of the simulation.

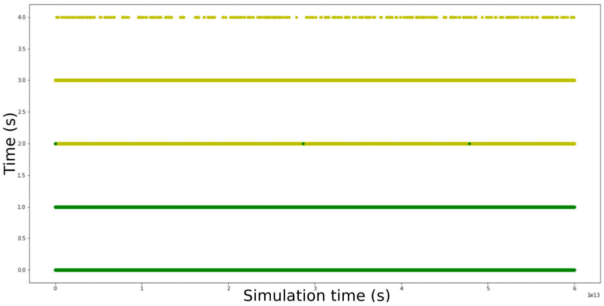

FIGURE 48. Queue behavior for CBS with high congestion.

TABLE 50. Queue behavior for CBS with high congestion and critical traffic.

	simtimeRaw	value
count	5.400600e+05	540060
mean	3.000268e+13	0.444506
std	1.731704e+13	0.496911
min	2.343100e+08	0
25,00%	1.500660e+13	0
50,00%	3.000671e+13	0
75,00%	4.499267e+13	1
max	5.999985e+13	1

TABLE 51. Queue behavior for CBS with high congestion and random traffic.

	simtimeRaw	value
count	1.420130e+06	1.420130e+06
mean	3.000412e+13	7.936104e-01
std	1.732433e+13	7.434559e-01
min	2.893100e+08	0
25,00%	1.499132e+13	0
50,00%	3.001982e+13	1
75,00%	4.500833e+13	1
max	5.999998e+13	4

TABLE 52. Queue behavior for CBS with high congestion and multimedia traffic.

	simtimeRaw	value
count	1.933780e+05	193378
mean	3.001787e+13	0.379454
std	1.731930e+ 13	0.485252
min	3.853100e+08	0
25,00%	1.502441e+13	0
50,00%	3.003268e+13	0
75,00%	4.501863e+13	1
max	5.999963e+13	1

According to Tables 50, 51, 52 and 53, it can be seen that the queuing level continues to be a little higher for random traffic, taking into account its statistical generation rate. For the other traffics, a behavior without queuing is maintained, except for non-relevant traffic that reaches a maximum of 2 queuing packages.

Latency (see Figure 49 and Tables 54, 55, 56 and 57):

Figure 49 shows latency results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). Under this saturation level, it is observed that critical traffic is the one that achieves the lowest latency times; however, it is interesting to observe that

TABLE 53. **Queue behavior for CBS with high congestion and no prioritized traffic.**

	simtimeRaw	value
count	2.869190e+05	286919
mean	2.999984e+13	0.477246
std	1.731989e+13	0.499525
min	3.817100e+08	0
25,00%	1.499981e+13	0
50,00%	3.000135e+13	0
75,00%	4.499532e+13	1
max	5.999990e+ 13	2

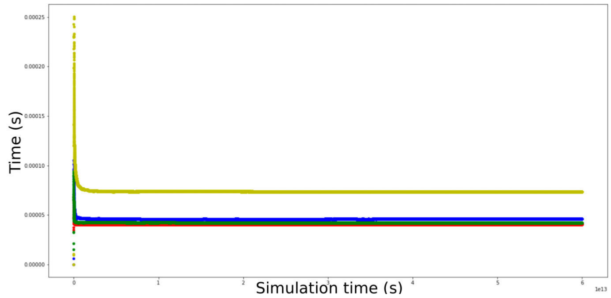

FIGURE 49. **Latency for CBS with high congestion.**

TABLE 54. **Latency for CBS with high congestion and critical traffic.**

	simtimeRaw	value
count	2.999990e+05	2.999990e+05
mean	2.999993e+13	4.065647e-05
std	1.732048e+13	1.789190e-07
min	2.786800e+08	0
25,00%	1.500000e+13	4.063652e-05
50,00%	2.999996e+13	4.065348e-05
75,00%	4.499987e+13	4.066894e-05
max	5.999979e+13	6.915693e-05

TABLE 55. **Latency for CBS with high congestion and random traffic.**

	simtimeRaw	value
count	7.501940e+05	7.501940e+05
mean	3.000433e+13	4.234000e-05
std	1.732254e+13	3.557534e-07
min	3.096800e+08	0
25,00%	1.499687e+13	4.2.29445e-05
50,00%	3.001772e+13	4.231513e-05
75,00%	4.500490e+13	4.233402e-05
max	5.999983e+13	9.431051e-05

random traffic has a higher latency than multimedia traffic and is not Relevant. This is explained due to the impact of critical traffic on attention times and that the random traffic generation rate is higher, thus increasing the number of packets that remain in the queue.

Although there is increased latency for random traffic, CBS traffic schedulers help ensure stable latency levels for critical traffic.

Jitter (see Figure 50):

Figure 50 shows jitter results considering critical, random, multimedia and no prioritized traffic (red, green, blue and yellow, respectively). Under this level of saturation, a wider distribution of jitter values is observed due to the times associated with the queuing of each of the traffics that are handled in the topology. Additionally, as there is a much higher channel utilization, the waiting time for a packet is higher to be transmitted on the output channel. As observed in Figure 50, there is an increase in the jitter distribution, which was expected due to the increase in the use of the channel.

TABLE 56. **Latency for CBS with high congestion and multimedia traffic.**

	simtimeRaw	value
count	1.199990e+05	119999
mean	2.999978e+13	0.000046
std	1.732044e+13	0.000001
min	5.467200e+08	0
25,00%	1.500004e+13	0.000046
50,00%	2.999974e+13	0.000046
75,00%	4.499949e+13	0.000046
max	5.999924e+13	0.000106

TABLE 57. **Latency for CBS with high congestion and no prioritized traffic.**

	simtimeRaw	value
count	1.499990e+05	149999
mean	2.999985e+13	0.000074
std	1.732045e+13	0.000003
min	8.289600e+08	0
25,00%	1.50001Oe+13	0.000074
50,00%	2.999992e+13	0.000074
75,00%	4.499968e+13	0.000074
max	5.999944e+13	0.000250

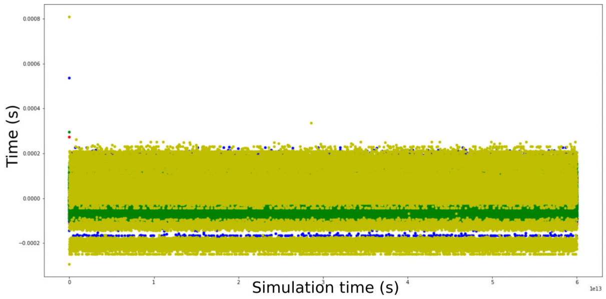

FIGURE 50. **Jitter for CBS with high congestion.**

All results presented in this section (results section V) are analyzed in depth in the next section (Analysis of results section VI).

VI. ANALYSIS OF RESULTS

The results obtained confirm the hypotheses about the positive impact of the use of TSN and EDGE Computing schemes in IIoT schemes in order to ensure the network parameters required by the deployed services. Next, a comparative analysis of the behavior of each one of the services in the different saturation scenarios is presented, starting from the services of higher priority to those of lower priority.

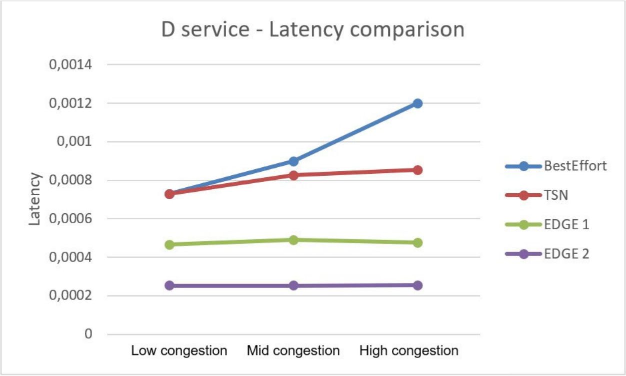

FIGURE 51. D service - latency comparison.

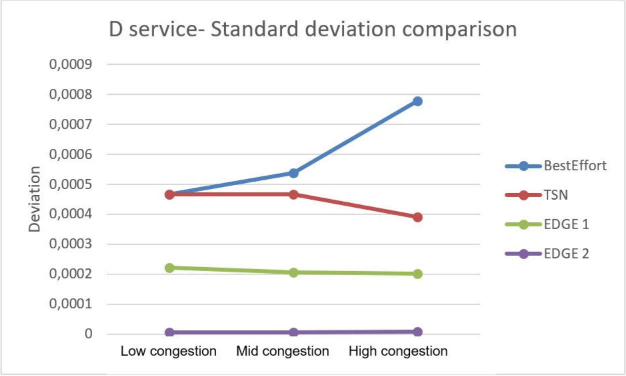

FIGURE 52. D service - standard deviation comparison.

A. D SERVICE

Service D was the one with the highest priority, and it simulated monitoring services with transmission of low-size packets with high frequency. In the TSN optimization processes, priority queue 6 was assigned for this service, and the queuing algorithms were configured through CBS (Credit Based Shaper), allowing that network devices could use this priority queue and, in this sense, avoiding longer queuing times due to service processes such as Strict Priority. Additionally, for this service, the deployment optimization process was performed using EDGE technologies in order to bring the service nodes closer to the IIoT devices. This model was performed in two stages: the first one generating 2 EDGE nodes for attending the IIoT, and the second stage, which involved the deployment of four attention nodes.

As can be seen in Figure 51, the implementation of TSN and EDGE was able to maintain stable latencies in the different congestion scenarios, which allows us to establish the effectiveness of these models to ensure the network parameters for critical services. The real impact on the behavior of the traffic in the topology depends largely on the ability to add EDGE attention nodes and their possible location, since this not only reduces the number of hops, but also can reduce the load on other elements of the topology.

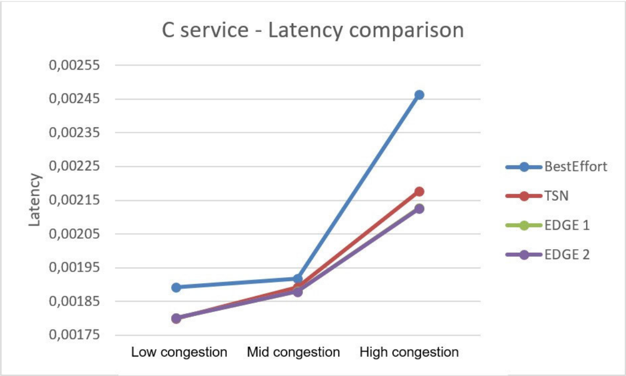

FIGURE 53. C service - latency comparison.

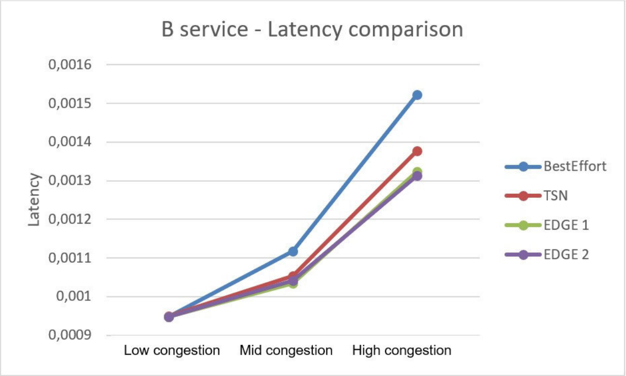

FIGURE 54. B service - latency comparison.

On the other hand, when observing Figure 52, we find that TSN services generate stability in the packets transmission, providing a reduction in the standard deviation of latency times, even when high levels of congestion are present. This is very important since this allows much more efficient utilization processes of the topology, by being able to maintain utilization levels above 50% without seeing a high impact on critical services.

B. C AND B SERVICES

Services C and B have an interesting effect on the results obtained, since, although they are not the highest priority services, they obtain benefits from the implementation of the TSN and EDGE schemes. For the queuing models, 5 and 4 priority queues were used for services C and D, respectively. These changes, combined with the positive effect of the change in the deployment topology due to the inclusion of EDGE nodes, achieve an effect of reducing the average latency compared to a Best-Effort scheme, as shown in Figures 53 and 54.

The effect of the reduction is more pronounced for service C, since it has a higher priority, thus achieving a latency reduction in queuing times. It is interesting to observe that there is not a considerable reduction between the average latency for EDGE 1 and EDGE 2 scenarios, which indicates

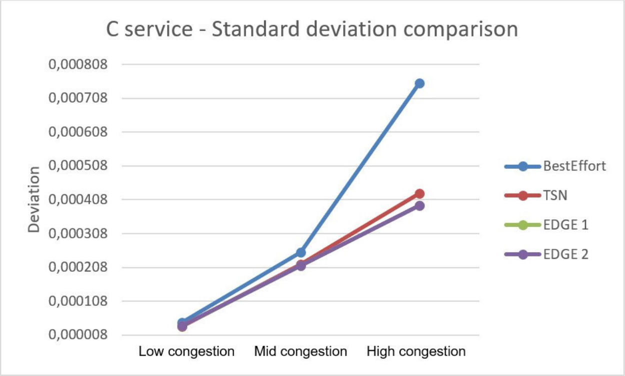

FIGURE 55. C service - standard deviation comparison.

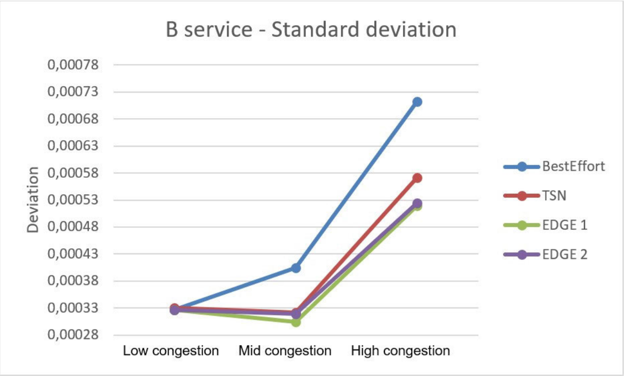

FIGURE 56. B service - standard deviation comparison.

that the increasing time is due to more IIoT devices are deployed, and that the highest priority traffic is not having an important impact to the other services in these scenarios.

As shown in Figures 55 and 56, there is a reduction in the time variability compared to the Best Effort model. However, for service C, an increase in the standard deviation of latency is observed as the traffic congestion increases in the network, which implies that the volume of traffic generated by the IIoT devices are responsible for queuing. This last situation generates variability in latency times, since very similar values are observed compared to the EDGE 1 and EDGE 2 scenarios, which eliminates the impact of service D traffic on the topology. On the other hand, service B behaves somewhat differently, since in the medium congestion scenario there is a reduction in latency variation. This may be due to stability in the volume of traffic generated, which maintains similar the latency values, although these are raised compared to the low-congestion scenario. Finally, it is observed that there is no significant difference for service B compared to the EDGE 1 and EDGE 2 scenarios, which means that with the optimization processes it was possible to reduce the impact on the topology bottleneck, which in this case are the links from SW05 and SW07 to SW08.

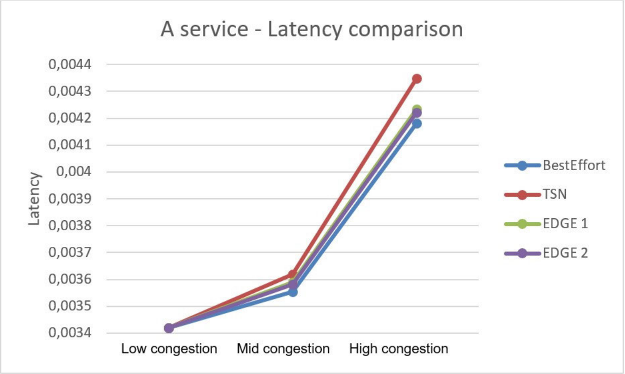

FIGURE 57. A service - latency comparison.

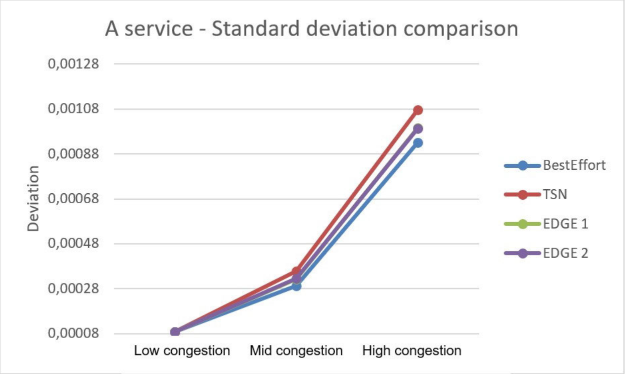

FIGURE 58. A service - standard deviation comparison.

C. A SERVICE

Service A, unlike the other services, had a negative impact with the implementation of the different optimization technologies, since, being the traffic with the lowest priority, it is the one that has the highest increases in queuing times due to the priority algorithms, attention queues and packet breaking.

As can be seen in Figure 57, the lowest average latency times were obtained with Best-Effort, since under this model all traffic is served under a FIFO scheme. Therefore, the impact of queuing is equitable among all services. Observing the other scenarios, we find that, for this service, the worst case occurs with the implementation of TSN, since it affects the queuing times, while the implementation of the EDGE 1 and 2 models reduces the average latency time by reducing or eliminating the impact of service D traffic on service A queuing times.

With respect to latency variability, as shown in Figure 58, we have the same effect, that is, the best scenario for service A is under Best-Effort and the worst scenario is with TSN. Maintaining the behavior observed in the other scenarios, we see that the impact on the latency times of the other services due to the implementation of the EDGE 1 and EDGE 2 models for service D, does not present a relevant variation, since with the EDGE 1 model, the reduction in the topology

TABLE 58. Latency analysis of the extended simulation.

Latency	Low congestion		Mid congestion		High congestion	
Traffic type	Best effort	CBS	Best effort	CBS	Best effort	CBS CBS
Critical	6e-05	2e-05	6e-05	6e-05	2e-04	4e-05
Random	6e-05	3e-05	8e-05	5e-05	2e-04	4e-05
Multimedia	7e-05	3e-05	8e-05	5e-05	2e-04	5e-05
Not relevant	6e-05	8e-05	7e-05	7e-05	2e-04	7e-05

TABLE 59. Jitter analysis of the extended simulation.

Jitter	Low congestion		Mid congestion		High congestion	
Traffic type	Best effort	CBS	Best effort	CBS	Best effort	CBS CBS
Critical	2e-08	7e-10	1e-08	4e-10	7e-09	6e-10
Random	7e-09	2e-10	4e-09	1e-10	3e-09	2e-10
Multimedia	5e-08	7e-09	4e-08	3e-09	2e-08	2e-09
Not relevant	4e-08	3e-09	2e-08	2e-09	2e-08	1e-09

bottleneck is generated, so the real effect between these two scenarios for services C, B and, in particular A, is almost negligible.

D. EXTENDED SIMULATION ANALYSIS

When comparing the data between the traffic scheduling schemes, an improvement is observed in all scenarios, particularly for high-priority traffic, although there are improvements in all types of traffic by managing to have a much more orderly and efficient queue management. much more efficient management processes.

As can be seen in Table 58, there is a significant improvement in the latency times obtained in the Best Effort models compared to the service times using devices with TSN capabilities in the topology. It is important to note that within the TSN topology, the latency times maintain the expected behavior where the traffic with the highest priority is the one with the lowest latency time compared to the results in Best Effort, where the attention time will depend on the level of general queuing and the traffic that is currently transmitting.

When evaluating Jitter (see Table 59), results similar to those obtained in latency are obtained, where the variability of the arrival times of the packets for the different traffics is lower in the model in which TSN was implemented. This behavior is expected, since by obtaining shorter attention times and distributed queue management by reducing the number of queued packets that must wait for transmission, it is possible to reduce the variability in arrival times at the destination application.

An additional important element to take into account in the use of the TSN topology was the configuration of the IdleSlope values, since this element is the one that provides versatility to program the attention spaces of each of the queues. If the value is very high, that queue will have a lot of priority and, therefore, it will be able to transmit much more compared to queues of less relevance, but if, on the contrary, the value is low, it can allow the queuing levels to be managed for the different types of traffic that are processed in the infrastructure.

In terms of interoperability and scalability of the solution, our proposal was very focused on the deployment of EDGE applications, that is, how TSN solutions were handled at layer 2. This scheme is transparent to any messaging that is used within the IIoT environment, whether it is MQTT or any other protocol. This allows interoperability between the deployment of TSN and EDGE models. On the other hand, regarding scalability, since TSN manages resource reservation processes, the scalability of the model is dictated by latency constraints and the number of devices that will be modeled within the topology. The TSN model can be adjusted by increasing network capabilities or distributing traffic in the topology, for example, by having multiple servers as endpoints for EDGE services to ensure compliance with latency and jitter requirements. In this sense, the traffic flow would be supported by the traffic reservation schemes implemented by TSN.

Finally, TSN provides very powerful tools that allow ensuring network parameters at the level of latency, jitter, and reliability for critical applications, such as all control systems that are currently being used in IIoT devices. However, an important aspect to continue exploring in the use of TSN capabilities, such as CBS traffic schedulers, is to be able to adjust the behavior of the network in such a way that priority traffic requirements are met and at the same time mitigate the effect of this on less relevant traffic, increasing the efficient usage that can be made of the communication channels.

VII. CONCLUSION

Through the results presented above, we obtained the main following conclusions:

- TSN provides multiple elements to ensure network parameters for critical services in IIoT environments. However, it is necessary to consider the negative impact that it can have on low priority services in the topology.
- EDGE computing models allow us to adjust the deployment topology, generating optimization processes, reducing the possible bottlenecks in the topology, and reducing the latency by diminishing the number of hops in from origin-destination routes.
- The topology design has a fundamental role in the amount of resources and their level (amount of switches, link speed, queuing capacity), since, although TSN and EDGE provide schemes to achieve optimization processes, the actual impact depends on the proposed topology.
- According to the deployment schemes, it is highly important to take into account the requirements of the IIoT service applications, due to this can facilitate or block EDGE computing schemes, since by requiring

fully centralized schemes, the network flexibility is considerably reduced.

- According to the messaging scheme, the way in which the IIoT devices perform the transmission of information to the attention nodes, allows us to determine deployment strategies subjected to this, that is, a direct scheme for TCP/UDP messaging, or subscription frameworks such as MQTT.
- EDGE optimization processes are delimited by the restrictions of the services to be deployed and, at the same time, by the costs related to the EDGE nodes implementation, since in the best scenario, EDGE nodes will be deployed in edge switches, reducing to the minimum the latency times. However, this scenario substantially increases the topology costs, whereby is necessary an optimization process to adjust the amount of EDGE nodes, in such a way that the network requirements for the services are accomplished and the total cost of the topology is minimized.
- Using TSN and EDGE technologies as complementary models, allows us to design highly efficient topologies with traffic assurance schemes, providing reliability for IIoT services that are increasingly relevant in Industry 4.0 and, also, that have more requirements.

REFERENCES

[1] (2019). *Value Creation and Capture: Implications for Developing Countries*. Digital Economy Report. United Nations Conference on Trade and Development. [Online]. Available: https://unctad.org/publication/digital-economy-report-2019

[2] (2016). *Why IIoT ES Different From Iot*. Intellinium. [Online]. Available: https://intellinium.io/why-iiot-is-different-from-iot/

[3] A. Nasrallah et al., "Ultra-low latency (ULL) networks: The IEEE TSN and IETF DetNet standards and related 5G Ull research," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 1, pp. 88–145. [Online]. Available: https://ieeexplore.ieee.org/document/8458130

[4] F. H. P. Fitzek, "Quality of service support in wireless multi-code CDMA systems," Ph.D. dissertation, Aalborg Univ., Denmark, U.K., 2002. [Online]. Available: https://www.researchgate.net/publication/227943040_Quality_of_Service_Support_in_Wireless_Multi-Code_CDMA_Systems

[5] S. Al-Sarawi, M. Anbar, R. Abdullah, and A. B. Al Hawari, "Internet of Things market analysis forecasts, 2020–2030," in *Proc. 4th World Conf. Smart Trends Syst., Secur. Sustainability (WorldS4)*, Jul. 2020, pp. 2020–2030.

[6] J. Farkas, L. L. Bello, and C. Gunther, "Time-sensitive networking standards," *IEEE Commun. Standards Mag.*, vol. 2, no. 2, pp. 20–21, Jun. 2018, doi: 10.1109/mcomstd.2018.8412457.

[7] N. Finn, "Introduction to time-sensitive networking," *IEEE Commun. Standards Mag.*, vol. 2, no. 2, pp. 22–28, Jun. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8412458

[8] J. L. Messenger, "Time-sensitive networking: An introduction," *IEEE Commun. Standards Mag.*, vol. 2, no. 2, pp. 29–33, Jun. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8412459

[9] B. R. Tunison, "The future of industrial alcohols," *J. Ind. Eng. Chem.*, vol. 12, no. 4, pp. 370–376, Apr. 1920, doi: 10.1021/ie50124a022.

[10] L. Bassi, "Industry 4.0: Hope, hype or revolution?" in *Proc. IEEE 3rd Int. Forum Res. Technol. Soc. Ind.*, 2012, pp. 190–210. [Online]. Available: https://ieeexplore.ieee.org/document/8065927

[11] B. Varghese, C. Reano, and F. Silla, "Accelerator virtualization in fog computing: Moving from the cloud to the edge," *IEEE Cloud Comput.*, vol. 5, no. 6, pp. 28–37, Nov. 2018, doi: 10.1109/MCC.2018.064181118.

[12] R. Shahzadi, A. Niaz, M. Ali, M. Naeem, J. J. P. C. Rodrigues, and S. M. Anwar, "Three tier fog networks: Enabling IoT/5G for latency sensitive applications," *China Commun.*, vol. 16, pp. 1–11, Dec. 2015. [Online]. Available: https://ieeexplore.ieee.org/document/8673718

[13] P. Pop, M. L. Raagaard, M. Gutierrez, and W. Steiner, "Enabling fog computing for industrial automation through time-sensitive networking (TSN)," *IEEE Commun. Standards Mag.*, vol. 2, no. 2, pp. 55–61, Jun. 2018, doi: 10.1109/mcomstd.2018.1700057.

[14] G.-J. Cheng, L.-T. Liu, X.-J. Qiang, and Y. Liu, "Industry 4.0 development and application of intelligent manufacturing," in *Proc. Int. Conf. Inf. Syst. Artif. Intell. (ISAI)*, Jun. 2016, pp. 407–410, doi: 10.1109/ISAI.2016.0092.

[15] I. Farris, L. Militano, M. Nitti, L. Atzori, and A. Iera, "Federated edge-assisted mobile clouds for service provisioning in heterogeneous IoT environments," in *Proc. IEEE World Forum Internet Things*, Jan. 2016, pp. 591–596. [Online]. Available: https://ieeexplore.ieee.org/document/7389120

[16] N. Chen, Y. Yang, J. Li, and T. Zhang, "A fog-based service enablement architecture for cross-domain IoT applications," in *Proc. IEEE Fog World Congr.*, Dec. 2018, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/document/8368533

[17] C. Chang, S. N. Srirama, and R. Buyya, "Indie fog: An efficient fog-computing infrastructure for the Internet of Things," *Computer*, vol. 50, no. 9, pp. 92–98, 2017, doi: 10.1109/MC.2017.3571049.

[18] *IEEE Standard for Local and Metropolitan Area Networks—Virtual Bridged Local Area Networks Amendment 12: Forwarding and Queuing Enhancements for Time-Sensitive Streams*, Standard 802.1Qav-2009, 2009. [Online]. Available: https://ieeexplore.ieee.org/document/5375704

[19] *IEEE 802.1Q: Media Access Control Bridges and Virtual Bridged Local Area Networks*. Accessed: Apr. 5, 2023. [Online]. Available: https://www.ieee802.org/802_tutorials/2013-03/8021-IETF-tutorial-final.pdf

[20] Q. Li, D. Li, and X. Sun, "Key technologies of time-sensitive networking and its application," in *Proc. Mater. Sci. Eng. Conf.*, 2005, pp. 101–155. [Online]. Available: https://iopscience.iop.org/article/10.1088/1757-899X/782/4/042015

[21] M. H. Farzaneh, "A modeling framework to facilitate schedule synthesis of time-sensitive networking," Ph.D. Dissertation, Tech. Univ. Munich, München, Germany, 2019.

[22] J. Falk, D. Hellmanns, B. Carabelli, N. Nayak, F. Durr, S. Kehrer, and K. Rothermel, "NeSTiNg: Simulating IEEE time-sensitive networking (TSN) in OMNeT++," in *Proc. Int. Conf. Networked Syst. (NetSys)*, Mar. 2019, pp 1–18, doi: 10.1109/NetSys.2019.8854500.

[23] H. R. A. Ameer and H. M. Hasan, "Enhanced MQTT protocol by smart gateway," *Iraqi J. Comput., Commun., Control Syst. Eng.*, vol. 3, pp. 53–67, Jan. 2021, doi: 10.33103/uot.ijccce.20.1.6.

[24] J. Perez, "MQTT performance analysis with OMNeT++ institut eurcom," M.S. thesis, IBM Zurich Res. Lab., Rüschlikon, Switzerland, 2005. [Online]. Available: https://doc.omnetpp.org/publications/Perez2005.pdf

[25] K. S. Kiangala and Z. Wang, "An effective communication prototype for time-critical IIoT manufacturing factories using zero-loss redundancy protocols, time-sensitive networking, and edge-computing in an industry 4.0 environment," *Processes*, vol. 9, no. 11, p. 2084, Nov. 2021, doi: 10.3390/pr9112084.

[26] A. Naouri, H. Wu, N. A. Nouri, S. Dhelim, and H. Ning, "A novel framework for mobile-edge computing by optimizing task offloading," *IEEE Internet Things J.*, vol. 8, no. 16, pp. 13065–13076, Aug. 2021, doi: 10.1109/JIOT.2021.3064225.

[27] T. Zhu, S. Dhelim, Z. Zhou, S. Yang, and H. Ning, "An architecture for aggregating information from distributed data nodes for industrial Internet of Things," *Comput. Electr. Eng.*, vol. 58, pp. 337–349, Feb. 2017.

DAVID L. GOMEZ received the degree in system engineering and the master's degree in systems and computing engineering from Universidad de Los Andes, Bogotá, Colombia, in 2011 and 2021, respectively. He is currently with the telecommunication industry in Bogotá. His main research interests include TSN networks and simulation.

GERMÁN A. MONTOYA received the degree in electronics engineering from Universidad Pontificia Bolivariana, Medellín, Colombia, in 2006, and the Ph.D. degree in engineering from Universidad de Los Andes, Bogotá, Colombia, in 2017. He is currently a Postdoctoral Assistant with the Department of Systems and Computing Engineering, Universidad de los Andes. His main research interests include mathematical optimization, simulation, and wireless sensor networks.

CARLOS LOZANO-GARZON (Senior Member, IEEE) received the degree in system engineering from Universidad Nacional de Colombia, Bogotá, Colombia, in 2002, and the Ph.D. degree from Universidad de los Andes, Colombia, and Universitat de Girona, Spain, in 2017. He was a Professor with Universidad Nacional de Colombia, among others, and as a Visiting Professor with Universidad Nacional de Asuncin, Paraguay. He is currently an Assistant Professor with the Department of Systems and Computing Engineering, Universidad de los Andes. His research interests include network optimization from a multi-criteria perspective, smart mobility: vehicle networks and intelligent transport systems, the Internet of Things, and network security.

YEZID DONOSO (Senior Member, IEEE) received the degree in system and computing engineering from Universidad del Norte, Barranquilla, Colombia, in 1996, the M.S. degree from Universidad de los Andes, Bogotá, Colombia, in 1998, and the D.E.A. and Ph.D. (cum laude) degrees in information technology from the University of Girona, Girona, Spain, in 2002 and 2005, respectively. He is currently a Full Professor with Universidad de los Andes and the Director of the Department of Systems and Computing Engineering. He is the coauthor of several books, including *Multi-Objective Optimization in Computer Networks Using Metaheuristics* and *Network Design for IP Convergence*. He has more than 240 papers published between journals and conferences. He has been a Senior Member of the IEEE Computer Society Distinguished Visitors Program (DVP), since 2005. He has several national and international awards and medals. From 2013 to 2014, he was the IEEE Colombia Section Chair.

• • •