Reducing Pessimism in Response Time Analysis of AVB Traffic in TSN

Daniel Bujosa*, Julian Proenza†, Alessandro V. Papadopoulos*, Thomas Nolte*, Mohammad Ashjaei*
* *Mälardalen University*, Västerås, Sweden
daniel.bujosa.mateu@mdu.se, alessandro.papadopoulos@mdu.se, thomas.nolte@mdu.se, mohammad.ashjaei@mdu.se
† University of the Balearic Islands, Palma, Spain
julian.proenza@uib.es

***Abstract*—Time-Sensitive Networking (TSN) is a set of standards with significant industrial impact potential, primarily due to its ability to integrate multiple traffic types with different requirements, offering great network flexibility. Among these traffic types, Audio-Video Bridging (AVB) stands out for its real-time guarantees and dynamic scheduling. In order to guarantee a specific set of AVB frames meet their timing requirements a Worst-Case Response Time Analysis (WCRTA) is essential. Unfortunately, current WCRTAs are often overly conservative, failing to guarantee schedulability for TSN systems operating even under low bandwidth conditions. This limits the practical usefulness of these analyses. Since TSN utilizes a multi-hop architecture, most WCRTAs analyze each link independently and then add the contributions. This compartmental analysis introduces pessimism, particularly when calculating the interference caused by other AVB frames with the same priority as the frame under analysis. In this paper, we address this issue by refining the Same-Priority Interference (SPI) calculation, leading to a significant improvement in the schedulability of WCRTAs and, consequently, the overall efficiency of TSN networks.**

***Index Terms*—Worst-Case Response Time Analysis, Time-Sensitive Networking, Audio-Video Bridging**

I. Introduction

Time-Sensitive Networking (TSN) has become one of the most relevant sets of standards for the industry. Among its many features, TSN enables Ethernet to integrate diverse traffic types with different characteristics and requirements simultaneously, providing significant network flexibility. This flexibility is achieved through TSN's three traffic classes: Scheduled Traffic (ST), Audio-Video Bridging (AVB) traffic, and Best-Effort (BE) traffic. AVB traffic is particularly valued for its real-time guarantees and adaptability. This is made possible by mechanisms such as the Credit-Based Shaper (CBS) and the Stream Reservation Protocol (SRP) described in the IEEE Std 802.1Q [1], which support online scheduling through bandwidth reservations and dynamic traffic configuration. However, while ST traffic is scheduled offline and guarantees its time requirements by construction, and BE traffic does not have real-time guarantees, to ensure that AVB traffic meets its timing requirements a Worst-Case Response Time (WCRT) Analysis (WCRTA) is required.

Since TSN operates on a multi-hop architecture, most existing WCRTAs analyze each link independently and aggregate the contributions along the path from source to destination. This approach of compartmental link analysis introduces unnecessary pessimism, especially when calculating the interference caused by same-priority frames, i.e. the interference caused by other AVB frames with the same priority as the frame under analysis. In this paper, we improve the Same-Priority Interference (SPI) calculation by considering the interactions between the input and output links of a switch. This refinement leads to a tighter analysis with reduced pessimism, resulting in improved schedulability and greater efficiency in TSN.

Contributions: In this paper, we focus on analyses that account for ST traffic, with and without preemption of non-ST traffic (i.e., AVB and BE), as this configuration is the most commonly used. However, any analysis that calculates the WCRT from source to destination can benefit from the enhanced STI calculation presented here. Therefore, we limit our scope to analyses that evaluate the WCRT across the entire path from source to destination. Specifically, we apply our solution to the analysis based on eligible intervals [2], as it is one of the most recent and has demonstrated significantly lower pessimism compared to earlier approaches, such as the WCRTA based on busy period [3] and the one based on Network Calculus [4], which meet the criteria for being considered in this paper. To the best of our knowledge, these are the only three AVB WCRTAs in the literature that consider ST interference and preemption according to the standards and compute the WCRT from the source to the destination. The concrete contributions in this paper are as follows:

- First, we identify a source of pessimism arising from the independent analysis of the output ports and their links, without considering the input links of the switch.
- Next, we extend the analysis to consider the input and output links at the switches, leading to a refined calculation of the SPI.
- Finally, we compare the improved SPI computation with results obtained using the WCRTAs that meet the properties considered in this paper, i.e. ST interference, both with and without preemption, and source-to-destination results.

Outline The paper is organized as follows. Section II describes the different mechanisms and features of TSN considered in this paper and Section III reviews the related works. Then, Section IV defines the system model, while Section V summarizes the WCRTA based on the eligible interval, and Section VI presents the pessimism problem identified in the previous analysis. Section VII extends the analysis through an improved SPI calculation. Finally, Section VIII presents the experimental setup, Section IX shows the results, and Section X concludes the paper.

II. TSN's Shapping Mechanisms and Features

In TSN, frames exchanged between end-stations are transmitted over paths composed of Ethernet links and TSN switches. The output ports of these switches and end-stations incorporate traffic-shaping mechanisms that support the various TSN traffic types. Specifically, each output port on both TSN end-stations and switches is equipped with up to 8 First-In-First-Out (FIFO) queues, each corresponding to one of the 8 priority levels available in TSN. Depending on the queue configuration and the integrated traffic-shaping mechanisms, each queue is assigned to handle one of the three TSN traffic types (ST, AVB, or BE). Each traffic type can utilize one or more of the available priorities, and thus, queues, resulting in multiple ST priorities and potentially multiple AVB and BE classes. For instance, the network can be configured to include three AVB classes (AVB class A, B, and C), each class having a lower priority than the previous one.

Fig. 1 illustrates a simplified 4-queue output port, comprising one ST queue, two AVB queues (Class A and B), and one BE queue. The following sections will provide a more detailed explanation of each TSN traffic shaper and their respective configurations.

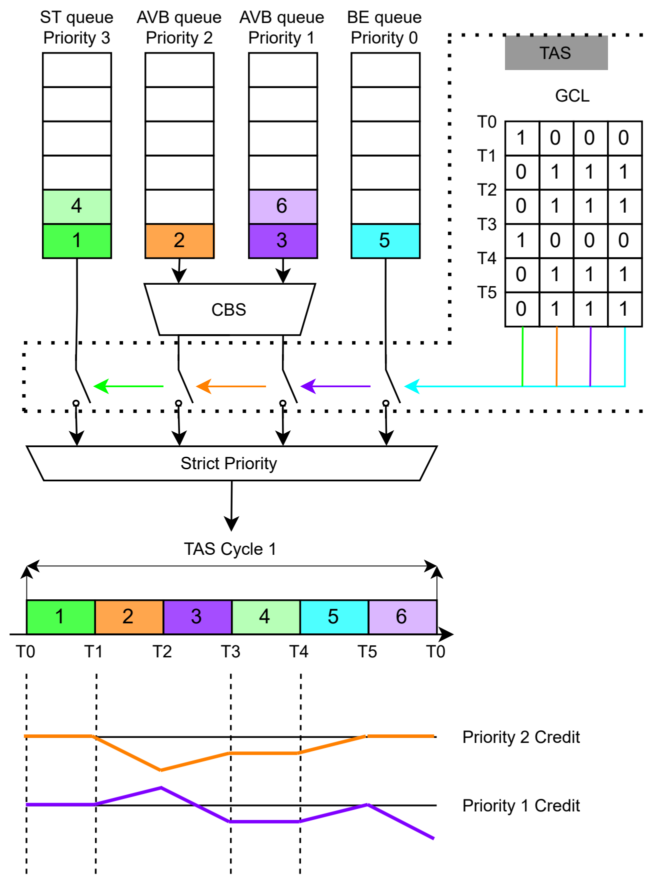


Fig. 1: Example of TSN output port.

A. Time-Aware Shaper

ST traffic is scheduled offline, ensuring fully deterministic transmission with zero jitter. To enforce this schedule, TSN utilizes the Time-Aware Shaper (TAS) mechanism, which assigns each queue a gate that can either be opened or closed. The gate's state is governed by the Gate Control List (GCL), a cyclically repeating list that specifies, at the nanosecond level, how long each gate remains open or closed. The cycle length is determined by the hyper-period of the ST frames transmitted through the corresponding output port, specifically the least common multiple of the periods of the ST frames passing through the gate. This characteristic of ST traffic is a limitation, as certain period combinations may cause an excessive increase in the GCL duration, which makes the use of AVB traffic interesting even for periodic frames.

When the gate is open, traffic from the queue can be transmitted, while a closed gate blocks transmission. The period during which the gate remains open is referred to as the *transmission window* . TSN also supports configuring ST traffic to preempt non-ST traffic, meaning that the transmission of a non-ST frame is temporarily halted and resumed after the ST frame has completed its transmission. Additionally, TAS can be configured to prevent the start of a new frame transmission if there is not enough time left until the gate closes to complete the transmission. However, this configuration will not be considered in our analysis, as it is rarely implemented in commercial switches.

For the transmission of an ST frame, the frame must reside in the designated queue, and its gate must remain closed until the scheduled transmission time. Before this time, all other queue gates are closed, during a certain time interval, to prevent interference. This time interval, when all gates are closed, is called the *guard band*. The size of the guard band depends on whether TSN is configured with or without preemption of non-ST traffic by ST frames. With preemption, the guard band must be at least the minimum preemptable frame size (124 bytes). Without preemption, the guard band must be larger than the maximum TSN frame size (1518 bytes). Omitting the guard band would introduce jitter in the transmission of ST frames. This jitter would be equivalent to the transmission time of the minimum preemptable frame size with preemption or the whole non-ST frame without preemption. Finally, the ST queue gate opens precisely at the scheduled transmission time, allowing the frame to be transmitted as scheduled. This mechanism is referred to as HOLD/RELEASE in the standards.

Fig. 1 provides an example with two ST frames (1 and 4). For simplicity, this example excludes guard bands and preemption. According to the GCL, these frames are scheduled for transmission at times T0 and T3, respectively. These transmission windows are marked by the dashed vertical lines in the figure. As shown, at T0, the ST gate is open while all other gates are closed, ensuring interference-free transmission of frame 1. Between T1 and T3, the ST gate is closed, and the other gates are open, allowing non-ST traffic to be transmitted. At T3, the ST gate is reopened, and the other gates are closed, enabling the transmission of frame 4 without interference. Finally, the ST gate closes again, and the non-ST gates open to allow the transmission of the remaining non-ST frames.

B. Credit-Based Shaper

While ST is transmitted on a precise, fixed schedule without jitter, its reliance on offline scheduling reduces its flexibility and limits its effectiveness in accommodating other types of traffic, such as event-triggered traffic. Additionally, the computational cost of obtaining a schedule rises considerably as the number of frames increases, potentially making scheduling infeasible or impractical in some scenarios. This is where AVB traffic proves advantageous, as it maintains real-time properties without being bound by the strict constraints of fixed scheduling.

To support real-time properties for AVB traffic, TSN employs the CBS mechanism. CBS restricts the maximum bandwidth percentage that a queue can utilize, ensuring that lower-priority queues receive a guaranteed minimum bandwidth allocation. In CBS, each AVB queue has a designated credit, which accumulates when an AVB frame is awaiting transmission or the credit is negative and is consumed while the queue is transmitting frames. Credit replenishment and consumption rates are defined by the terms *idleSlope* and *sendSlope*, respectively. A queue can transmit only when its credit is zero or positive, and its credit is frozen if the gate is closed.

In the example in Fig. 1 a single frame (frame 2) is allocated to the higher-priority AVB queue, while the lower-priority AVB queue holds two frames (frames 3 and 6), and the BE queue contains one frame (frame 5). At time T1, following the transmission of the ST frame, both AVB queues become eligible for transmission. The strict priority mechanism then selects the higher-priority AVB queue, resulting in the transmission of frame 2. During this transmission, the credit of the higher-priority AVB queue is consumed, while the credit of the lower-priority AVB queue is replenished as it awaits transmission.

At T2, the higher-priority AVB queue, now with negative credit, becomes ineligible for transmission. The lower-priority AVB queue is then selected, leading to the transmission of frame 3. Similar to T1, the credit of the active AVB queue is consumed during frame 3's transmission, while the credit for the higher-priority AVB queue is replenished.

At T3, another ST frame is transmitted, closing the gates of the AVB queues and freezing their credits. By T4, both AVB queues have negative credit, allowing the BE queue to transmit, even though a higher-priority AVB frame is awaiting transmission. Finally, at T5, the lower-priority AVB queue's credit reaches zero, permitting the transmission of the last frame (frame 6). Note that BE traffic lacks any timing requirements. Therefore, it can only transmit when its gate is open and no higher-priority frames are eligible for transmission.

III. Related Work

Since the introduction of the AVB standard in 2011, numerous works have tackled the challenge of analyzing its WCRT. These approaches are commonly classified into three main categories: busy period analysis, eligible interval analysis, and Network Calculus modeling.

Busy period analysis identifies the critical instant that produces the WCRT. Diemer et al. [5] pioneered AVB WCRT analysis based on this approach, although it was limited to a single output port and AVB queue, without accounting for ST interference. Later, in [6], they extended their work to handle two AVB queues (class A and class B) but remained constrained to a single output port and still omitted ST interference. In the more recent work of Lo Bello et al. [3], the authors present an AVB WCRTA based on busy period analysis that considers interference across multiple AVB classes and incorporates ST interference with and without preemption, as well as lower-priority frame blocking. This analysis also includes multi-hop calculations, extending the analysis from source to destination rather than limiting it to a single output port. However, it remains limited to two AVB classes (Class A and B).

The second approach, eligible interval analysis, defines the time interval in which each frame becomes eligible for transmission. The WCRT of a frame is calculated by analyzing the maximum interference it encounters over this interval. One of the earliest analyses of this type was introduced by Bordoloi et al. [7]. Cao et al. [8]–[10] later extended this work and formally introduced the term eligible interval, showing that the WCRT of an AVB frame depends solely on the maximum achievable credit of its AVB class, which is bounded and computable. This approach reduces WCRT pessimism compared to busy period analysis and applies to any number of AVB classes. However, it excludes ST interference and remains restricted to single output ports. Maxim et al. [11] extended the previous works to address ST interference without preemption.

The third approach employs Network Calculus, a mathematical framework for performance analysis in communication networks, to calculate the maximum delay, or WCRT, each frame can experience. Zhao et al. [4], [12] present a Network Calculus-based AVB WCRTA, allowing WCRT calculations from source to destination for multiple AVB queues with ST interference, both with and without preemption. Nevertheless, Network Calculus faces limitations when analyzing loop networks with circular dependencies and enforces equal bandwidth allocation for AVB classes across all links, reducing configuration flexibility. Other Network Calculus-based WCRTAs exist but were not included in this review due to limitations. For instance, [13] proposes a WCRTA tailored to a specific configuration with a single transmission window per traffic type and restrictions preventing concurrent AVB class A and B transmissions, limiting the CBS arbitration capabilities.

Finally, Bujosa et al. introduced a WCRTA in [2] that combines busy period and eligible interval analyses. This analysis, together with the proposals presented in [12], and [3], are the only ones that satisfy the requirements of this work, specifically providing source-to-destination WCRT calculations with the consideration for ST interference both with and without preemption. Due to its superior performance, we apply our improvement to the WCRTA proposed in [2] and compare the results with the previous WCRTA. Comparisons with the analysis based on busy periods are omitted, as the WCRTA combining busy periods and eligible intervals is an extension of it.

IV. System Model

This section provides network and system models required for the response time analysis.

A. Network model

Connections between any end-station and any TSN switch, as well as between two TSN switches, are established through *link*s represented by l. The TSN ports operate in full-duplex mode, meaning that input and output operations are isolated. As a result, reception and transmission on the same physical port do not interfere with each other. Consequently, each physical port corresponds to two links: one for transmission and another for reception. Additionally, the duration from the reception of a frame at a TSN switch port to its queuing into the output port queue is unique to each TSN switch, denoted as ϵ. For the purposes of this analysis, the link delay attributed to the wire and its physical characteristics is considered negligible, while the overall network bandwidth, represented by BW, remains consistent across all network links. Lastly, for AVB traffic of priority X on a link l, the credit replenishment rate (idleSlope) is designated as $\alpha^+_{X,l}$, the credit consumption rate (sendSlope) is denoted by $\alpha^-_{X,l}$, and the credit value for traffic in priority X on link l is represented as $CR_{X,l}$.

B. Traffic model

We adopt a real-time periodic model for all types of traffic in TSN networks. This model defines a stream as a sequence of frames that share common attributes, including source and destination addresses, periods, and deadlines. Accordingly, a collection of N streams is characterized as follows:

$$\Gamma = \{m_i(C_i, T_i, D_i, P_i, \mathcal{L}_i, \mathcal{O}_i)|i = 1, \ldots, N\} \qquad (1)$$

Within this model, C_i denotes the transmission time of a frame of the stream m_i, which is determined by the frame size and the network bandwidth. Note that the Ethernet frame header is included in the frame transmission time C_i, with the header transmission time indicated as v. This value is treated as constant and independent of frame type. For the ST, the guard band is included in C_i when necessary; notably, when multiple ST frames are transmitted sequentially without gaps between them, only the initial frame requires a guard band. Furthermore, T_i and D_i represent the period and relative deadline of the frames, respectively. We assume a constrained deadline model, meaning $D_i \leq T_i$. It is important to note that AVB traffic classes can be initiated either periodically or sporadically. In cases of sporadic transmissions, T_i signifies the minimum inter-arrival time, i.e. the shortest time between two frames of the stream m_i. The priority of a stream is indicated by $P_i \in \mathbb{P}$. The highest priority is assigned to ST, the lowest to BE, and all intermediate priorities are allocated to AVB. In this context, ST priority is denoted as P_{ST}, BE priority as P_{BE}, and AVB priorities as $\{P \in \mathbb{P} \mid P_{ST} > P > P_{BE}\}$. Additionally, $\mathbb{L}$ and $\mathbb{H}$ represent the sets of non-ST streams with lower and higher priority than m_i, respectively. Since a frame may traverse multiple links, the set of n links that m_i passes through is specified by $\mathcal{L}_i = \{\mathcal{L}_i(0), \ldots, \mathcal{L}_i(n)\}$.

Offsets are utilized to fit ST streams into the transmission schedule. The offset for each ST frame is defined per link, and the collection of offsets for all links traversed by m_i is given by $\mathcal{O}_i$, for example, $\mathcal{O}_i = \{O^l_i\}$. We assume that the offsets are predetermined, as the scheduling of ST streams is beyond the scope of this paper. Previous studies, such as [14], have already addressed this topic. It is important to note that AVB streams do not have defined offsets; thus, for AVB frames, $\mathcal{O}_i$ is represented as an empty set, i.e., $\mathcal{O}_i = \{\emptyset\}$.

In this model, we consider the TSN configuration where ST frames can preempt non-ST frames. This means that some frames may experience interruptions during transmission, which will be resumed later. Additionally, ST frames that arrive at a switch will remain in the output queue until their transmission window is activated. Specifically, traffic classes can be categorized as *express* or *preemptable*. In our system model, ST streams are classified as express, while all other classes are preemptable. This implies that the guard band can be set to 124 Bytes, representing the non-preemptable segment of any non-ST frame. This value indicates the maximum duration a preemptable frame (non-ST frame) can block an express frame (ST frame). This configuration is typical for TSN when all traffic types are considered, i.e. ST, AVB, and BE.

Lastly, when an ST frame preempts a non-ST frame, it resumes transmission with a new header. Thus, if a frame is preempted and divided into two segments, the second segment will also have a header, resulting in two headers needing to be accounted for. This has implications for the analysis of AVB traffic.

V. Worst-Case Response Time Analysis

This section offers an in-depth overview of the WCRTA based on eligible intervals [2]. As previously outlined, this WCRTA takes into account the interference from higher-priority streams, which encompasses ST interference with preemption and higher-priority AVB interference, same-priority interference, and blocking from lower-priority frames, including lower-priority AVB and BE. Furthermore, it considers the contributions of multi-hop behavior, providing the WCRT from the source to the destination.

A. WCRTA overview

The WCRTA based on busy periods and eligible intervals considers various sources of delay that an AVB frame of stream m_i may encounter on link l and adds them in a compositional way. First, since all queues in TSN operate on a FIFO basis, the transmission of a frame is contingent upon the preceding frames in the same queue. Consequently, the source of delay arises from interference with same-priority traffic, represented as SPI_i^l. Secondly, higher-priority AVB frames can interfere with lower-priority AVB frames, indicated by HPI_i^l, although this interference is somewhat limited due to the CBS. Furthermore, since AVB classes are non-preemptive, if frames from lower-priority classes are currently being transmitted, the AVB frame must wait for the entire transmission duration. This results in blocking caused by lower-priority traffic, denoted as LPI_i^l. Finally, the ST, via the TAS gates and according to the established schedule, can block the transmission of any AVB or BE queue. Given that the schedule is not uniform, it is essential to identify this interference, denoted as $STI_{i,c[k]}^l$, for all critical instant candidates $I_c^l[k]$ which correspond to the beginning of each transmission window of each frame k of each ST stream c on link l. In summary, to compute the WCRT of any AVB frames in the critical instant candidate $I_c^l[k]$, one must consider the contributions from ST, higher-priority AVB traffic, same-priority AVB traffic, and lower-priority traffic (both lower-priority AVB and BE). Thus, for a frame of stream m_i traversing link l at the critical instant candidate $I_c^l[k]$, the $WCRT_{i,c[k]}^l$ is expressed as:

$$WCRT_{i,c[k]}^l = STI_{i,c[k]}^l + HPI_i^l + SPI_i^l + LPI_i^l + C_i \quad (2)$$

Once the WCRT for a frame of m_i in link l is determined for each potential critical instant candidate $I_c^l[k]$, the WCRT of a frame of stream m_i in link l is established by identifying the maximum response time across all critical instant candidates. This means that the interference for each critical instant candidate must be assessed, and the highest value should be chosen as follows:

$$WCRT_i^l = \max_{\forall m_c, \forall k} \{ WCRT_{i,c[k]}^l \} \quad (3)$$

Finally, since the analysis is compositional, where a frame is buffered as it passes through each hop, we sum the $WCRT_i^l$ for each link along the path from the source to the destination of the frame m_i and add the ϵ factor for each switch crossed. This factor accounts for the delay incurred by the frame from its reception at the switch's input port until it is queued in the TSN output port. The overall WCRT for frame m_i is computed as follows:

$$WCRT_i = \sum_{l=1,\dots,|\mathcal{L}_i|} WCRT_i^l + (|\mathcal{L}_i| - 1) \times \epsilon \quad (4)$$

In the subsequent subsections, we will discuss the worst-case scenarios for each of the components that contribute to the WCRT for an AVB frame m_i on link l.

B. Same Priority Interference

As outlined in [2], the delay experienced by an AVB frame of stream m_i on link l, when subjected solely to interference from same-priority traffic, i.e., $sp(m_i) = \{m_j | P_j = P_i, j \neq i\}$, is determined based on a basic FIFO scheduling approach. However, due to the behavior of the CBS, it is essential to account not only for the transmission time of each same-priority frame C_j, but also for the time needed to recover the credit consumed by those interfering frames. When only same-priority traffic is involved, the credit level cannot be bigger than 0, as that would imply interference or blocking from other traffic classes. Thus, same-priority frames can only be transmitted when the credit reaches 0. These frames will consume $C_j \times \alpha_{P_i,l}^-$ credit, which needs to be replenished over a duration of $C_j \times \frac{\alpha_{P_i,l}^-}{\alpha_{P_i,l}^+}$ for the credit to replenish to 0, allowing for the transmission of the next frame in the queue. Consequently, the total interference is the sum of the transmission time of the interfering frames and the time required to restore the credit consumed by each of these frames. In the worst-case scenario, a frame may be interfered with by all same-priority streams. However, given that the analysis operates under a deadline-constrained model, only one frame of each same-priority stream in the FIFO queue can interfere with the frame under analysis if all frames satisfy their deadlines, as discussed in the context of the Controller Area Network (CAN) [15]. In this way, similar to most analyses with constrained deadlines, the results lack reliability if any frame misses its deadline according to its WCRT. Therefore, the interference from same-priority frames on m_i of class P_i on link l is computed using Eq. (5).

$$SPI_i^l = \sum_{\substack{\forall m_j \in sp(m_i), i \neq j \\ \wedge l \in \mathcal{L}_j}} C_j \times \left(1 + \frac{\alpha_{P_i,l}^-}{\alpha_{P_i,l}^+}\right) \quad (5)$$

C. Higher-Priority AVB Interference and Lower-Priority Blocking

While higher-priority AVB interference and lower-priority blocking represent distinct contributions, the authors in [10] demonstrated that these delay contributions correspond to the time required to achieve the maximum credit CR^{max} for the AVB class of the analyzed stream m_i of priority P_i, expressed as:

$$HPI_i^l + LPI_i^l = \frac{CR_{P_i,l}^{max}}{\alpha_{P_i,l}^+}. \quad (6)$$

The authors also established that $CR_{P_i,l}^{max}$, and consequently $HPI_i^l + LPI_i^l$, remains bounded provided that the total bandwidth assigned to P_i and all higher-priority queues $\mathbb{H} = \{H \in \mathbb{P} \,|\, ST > H > P_i\}$ does not surpass the available bandwidth, i.e.:

$$\sum_{\forall P \in \mathbb{H} \cup P_i} \alpha_{P,l}^+[\%] \leqslant BW[\%]. \quad (7)$$

Given these conditions, the non-ST interference $HPI_i^l + LPI_i^l$ experienced by a frame can be computed as follows:

$$\begin{aligned} HPI_i^l + LPI_i^l &= \frac{CR_{P_i,l}^{max}}{\alpha_{P_i,l}^{+}} \\ &= C_{\mathbb{L},l}^{max} \times \left(1 + \frac{\alpha_{\mathbb{H},l}^{+}}{\alpha_{\mathbb{H},l}^{-}}\right) - \frac{CR_{\mathbb{H},l}^{min}}{\alpha_{\mathbb{H},l}^{-}} \end{aligned} \tag{8}$$

where $C_{\mathbb{L},l}^{max}$ represents the size of the largest frame from all lower-priority queues $\mathbb{L} = \{L \in \mathbb{P} \mid L < P_i\}$ and $CR_{\mathbb{H},l}^{min}$ is the minimum value that the combined credit of the highest priority queues can achieve on link l. This latter value is computed recursively as follows:

$$\begin{aligned} &CR_{\mathbb{H}=\{H_1,\ldots,H_n\},l}^{min} = \\ &\quad -\max(\alpha_{\mathbb{H},l}^{-} \times C_{H_1,l}^{max} - CR_{\mathbb{H}-H_1,l}^{min}, \\ &\quad\quad \ldots, \alpha_{\mathbb{H},l}^{-} \times C_{H_n,l}^{max} - CR_{\mathbb{H}-H_n,l}^{min}) \end{aligned} \tag{9}$$

D. Scheduled Traffic Interference

As proven in [3], the starting time of each ST transmission window within the hyper-period must be considered as a critical instant candidate. Since every link has its unique hyper-period denoted as Ω_l, and the hyper-period for a set of frames is determined by the least common multiple of their respective periods, the specific instances relevant for assessing the ST interference of the frame $m_j \in ST$ on link $l \in \mathcal{L}_j$ are defined as:

$$I_j^l = \{(k-1)T_j + O_j^l \ : k = 1, \ldots, n \ , \ n = \frac{\Omega_l}{T_j}\} \tag{10}$$

After identifying all potential critical instants throughout the hyper-period for link l, the phase difference between each ST frame in $m_j \in ST$ and each potential critical instant $I_c^l[k]$ is calculated. These phase differences represent the offsets that different ST frames m_j would exhibit if the beginning of the hyper-period coincided with the critical instant candidate of frame k of stream $m_c \in ST$. For further details and supporting proofs, please refer to [16].

$$\Phi_{jc[k]}^l = (O_j^l - I_c^l[k]) \bmod T_j \tag{11}$$

Finally, for every critical instant candidate $I_c^l[k]$, the ST interference experienced by an AVB frame over time t is expressed as:

$$W_{c[k]}^l(t) = \sum_{\forall j \in ST \ \wedge l \in \mathcal{L}_j} \left(\left\lfloor \frac{\Phi_{jc[k]}^l}{T_j} \right\rfloor + \left\lceil \frac{t - \Phi_{jc[k]}^l}{T_j} \right\rceil \right) C_j \tag{12}$$

Additionally, AVB traffic may be preempted by each interfering ST frame, leading to the transmission of additional headers. Therefore, for every preemption caused by an ST frame, the added interference attributed to the header size v must be accounted for. Furthermore, these additional headers will consume credit that requires replenishment. Depending on whether the preemption affects a same-priority frame or a higher-/lower-priority frame, it will be weighted according to Eq. (5) or Eq. (8), respectively. In the worst-case scenario, the higher of the two cases will be selected, resulting in Eq. (13).

$$\begin{aligned} V_{c[k]}^l(t) = &\sum_{\forall j \in ST \ \wedge l \in \mathcal{L}_j} \left(\left\lfloor \frac{\Phi_{jc[k]}^l}{T_j} \right\rfloor + \left\lceil \frac{t - \Phi_{jc[k]}^l}{T_j} \right\rceil \right) v \\ &\times \left(1 + \max\left(\frac{\alpha_{P_i,l}^{-}}{\alpha_{P_i,l}^{+}}, \frac{\alpha_{\mathbb{H},l}^{+}}{\alpha_{\mathbb{H},l}^{-}}\right)\right) \end{aligned} \tag{13}$$

Consequently, the maximum ST interference that an AVB frame from stream m_i can encounter on link l at instant $I_c^l[k]$ over time t is computed in Eq. (14), which represents the total interference from ST and the additional headers resulting from preemption.

$$STI_{c[k]}^l(t) = W_{c[k]}^l(t) + V_{c[k]}^l(t) \tag{14}$$

In this manner, the response time of an AVB frame queued at the output port of link l during the critical instant candidate $I_c^l[k]$, represented as $WCRT_{i,c[k]}^{l,(x)}$, is iteratively calculated as follows:

$$\begin{aligned} WCRT_{i,c[k]}^{l,(x)} = &STI_{c[k]}^l\left(WCRT_{i,c[k]}^{l,(x-1)}\right) + \\ &HPI_i^l + SPI_i^l + LPI_i^l + C_i. \end{aligned} \tag{15}$$

The iteration starts with $WCRT_{i,c[k]}^{l,(0)} = HPI_i^l + SPI_i^l + LPI_i^l + C_i$ and concludes when $WCRT_{i,c[k]}^{l,(x)} = WCRT_{i,c[k]}^{l,(x-1)}$.

VI. Problem Formulation

In most WCRTAs and the analysis presented in Section V, a frame may be interfered with by one frame from each stream of the same priority under the assumption of constrained deadlines. However, this scenario can occur only at the transmitter's output port. In order to be interfered by all same-priority frames, those must arrive at the transmission queue simultaneously, just before the arrival of the frame under analysis. This situation is plausible at the talker's transmission queue, where applications may attempt to send frames concurrently, utilizing parallel resources. However, such simultaneous arrival of frames is unlikely at the switches, as frames are received sequentially through input ports. Consequently, frames of the same priority require reception times determined by the transmission rate non-null, leading to some frames being forwarded while the remaining frames are still being received. As a result, the maximum number of same-priority interferences will be less than or equal to the total number of frames of the same priority.

To analyze the problem, we will examine an extreme scenario involving a single switch with one input and output links. This switch receives and transmits traffic associated with a single AVB priority, i.e. all frames received are same-priority frames. Specifically, n streams of identical size C and period T are processed. Additionally, since all traffic is assigned the same priority, 100% of the bandwidth will be allocated to this priority, resulting in no credit recovery time. Fig. 2 shows the WCRT of a frame as calculated using the existing WCRTAs. In this figure, the horizontal lines represent

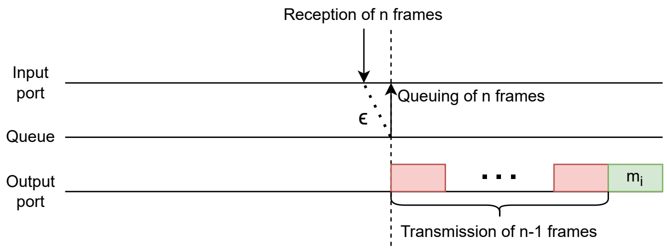


Fig. 2: WCRT of an AVB frame with a single input and output port and a single priority through traditional WCRTAs.

the evolution of the input and output ports, along with the AVB queue. According to conventional analysis, one frame from each same-priority stream, including the frame under analysis, arrives almost simultaneously through the input link (indicated by the downward arrow). Consequently, by the time the frame under analysis is queued (indicated by the vertical dashed line), the $n-1$ preceding frames will have already entered (indicated by the upward arrow), resulting in a delay of $n \times C$. However, a closer examination of the switch's actual behavior reveals a different outcome, as illustrated in Fig. 3. This figure demonstrates that the reception of the n frames takes a finite amount of time, which the output link utilizes to retransmit those same-priority frames. Consequently, when the frame under analysis reaches the AVB queue, no same-priority frames are available for interference, leading to a delay of C, i.e., a delay n times smaller than the obtained through traditional WCRTA.

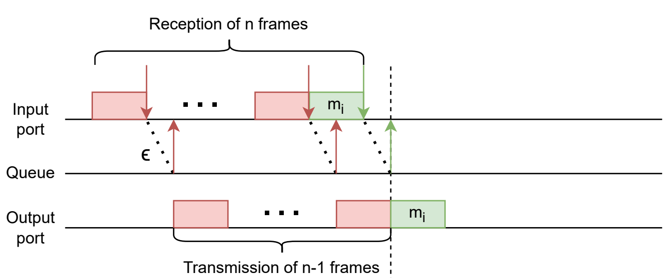


Fig. 3: WCRT of an AVB frame with a single input and output port and a single priority.

In the upcoming section, we will conduct a detailed analysis of the maximum same-priority interference that an AVB frame may encounter, reducing the pessimism of the calculation significantly.

VII. Proposed Solution

This section outlines the main contributions of our work, presenting the lemmas and proofs that lead to the calculation of the maximum SPI.

The key concept in calculating the maximum SPI is to determine the minimum time the queue of the frame under analysis can transmit same-priority traffic before the frame under analysis' arrival time to the queue.

We begin the proof by analyzing the case without blocking or interference from other traffic types. Subsequently, we examine the interactions with other traffic to determine how they affect the calculation of the maximum SPI.

A. SPI without Blocking nor Interference from other Classes

To calculate the minimum time a queue can transmit in the absence of blocking and interference, two key aspects must be considered. First, we need to determine the minimum time required to receive all same-priority frames, with the frame under analysis being the last to be received. In the worst-case scenario, this represents the maximum time the output queue can transmit same-priority frames that would typically interfere with the frame under analysis.

Definition 7.1: The Minimum Reception Time (MRT) is the shortest duration required to receive a set of frames across one or more communication links, considering transmission times and any dependencies between the frames, such as credit recovery in the case of AVB traffic.

Second, it is essential to calculate the minimum elapsed time between any two frames of the same stream. If the interval between a frame of a same-priority stream and its predecessor is very short, it could result in the transmission of the preceding frame occurring during the reception time of the same-priority frames. This situation limits the transmission of same-priority frames that could interfere with the frame under analysis. A detailed analysis of this scenario is provided below.

Definition 7.2: The Minimum Time Separation (MTS) is the shortest time interval that must elapse between the completion of a frame's transmission on a link and the start of the subsequent frame's transmission from the same stream on the same link.

Lemma 7.1: The minimum time necessary to receive the same priority frames as the frame under analysis (including the frame under analysis of stream m_i) through link $\{l' | \exists \mathcal{L}_j(x) = l' \;\&\; \mathcal{L}_j(x+1) = l\}$ that will be retransmitted by link l, referred to as Minimum Reception Time ($MRT_i^{l',l}$) is calculated as follows:

$$\zeta = \left(\sum_{\substack{\forall m_j \in sp(m_i) \\ \wedge l, l' \in \mathcal{L}_j}} C_j \right) - C_{P_i,l'}^{max} \tag{16}$$

$$MRT_i^{l',l} = \max\left(\zeta \times \left(1 + \frac{\alpha_{P_i,l'}^{-}}{\alpha_{P_i,l'}^{+}}\right) - \frac{CR_{P_i,l'}^{max}}{\alpha_{P_i,l'}^{+}}, \zeta \right) \tag{17}$$

where $C_{P_i,l'}^{max} = \max_{\forall m_j \in sp(m_i) \wedge l, l' \in \mathcal{L}_j} (C_j)$.

Proof. In the worst-case scenario, assuming no blocking nor interruptions, frames are received sequentially with an inter-frame interval that corresponds to the time required to recover the credit consumed during the transmission of each frame, as outlined in Section V and demonstrated in [8]–[10], i.e.: $\sum_{\substack{\forall m_j \in sp(m_i) \\ \wedge l, l' \in \mathcal{L}_j}} C_j \times \left(1 + \frac{\alpha_{P_i,l'}^{-}}{\alpha_{P_i,l'}^{+}}\right)$. Conversely, as illustrated in Fig. 3, the frame reception time for the transmission queue spans from the conclusion of the first frame's reception to the completion of the last frame's reception. Consequently, when calculating the total reception time, it is essential to exclude both the transmission time of the first frame and the credit recovery time of the last frame. In order to ensure a minimum frame reception time, the excluded times must be maximized. This is effectively equivalent to omitting the transmission and credit recovery time associated with the largest frame, denoted as $C^{max}_{P_i,l'}$, from the $MRT^{l',l}_i$ calculation. Additionally, the credit at the start of same-priority frame reception will be the maximum achievable by the queue, meaning that any credit already accumulated prior to transmission must be excluded from the credit recovery calculation, i.e. $-\frac{CR^{max}_{P_i,l'}}{\alpha^{+}_{P_i,l'}}$. However, this reduction in credit recovery time cannot result in a reception time shorter than the duration required to receive all same-priority frames, i.e. ζ. ■

In this regard, the minimum time necessary to receive all the same priority frames as the frame under analysis (including the frame under analysis of stream m_i) that will be forwarded by link l is calculated as follows:

$$MRT^l_i = \max(MRT^{l',l}_i) \tag{18}$$

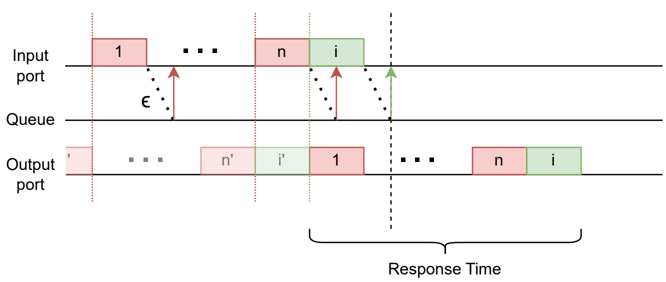


Fig. 4: WCRT of an AVB frame with a single input and output port and a single priority considering stream's previous frames.

During MRT^l_i, the link l can forward part of the same-priority traffic. However, a portion of this time might be used for transmitting previous frames of these same-priority streams. Figure 4 illustrates an extreme case where, just before the reception of each same-priority frame on the input link (frames $1..n, i$), a previous frame of the same stream is sent through the output link (frames $1'..n', i'$). This scenario consumes nearly all the MRT^l_i time in transmitting previous frames of the same-priority streams.

In this context, we must calculate the minimum temporal distance between two consecutive frames of the same stream. This calculation will determine whether a same-priority frame arriving at the queue before the frame under analysis could have a previous frame of the same stream transmitted during MRT^l_i. Specifically, we are interested in the temporal distance between the end of the transmission, including the recovery of the credit of a frame on link l, and the start of the transmission of the next frame of the same stream on the same link l.

Lemma 7.2: The Minimum Temporal Separation (MTS^l_j) between the end of the transmission, including the recovery of the credit of a frame of stream m_j on link l, and the start of the transmission of the next frame of the same stream on the same link l occurs when one frame experiences the WCRT across the set of links $\mathcal{L}^l_j = \mathcal{L}_j(0), \ldots, \mathcal{L}_j(x) = l$, and the subsequent frame experiences the Best-Case Response Time (BCRT). This separation is calculated as follows:

$$MTS^l_j = D_j + (|\mathcal{L}^l_j| - 1) \times C_j - C_j \times \frac{\alpha^{-}_{P_j,l}}{\alpha^{+}_{P_j,l}} - \sum_{\forall l'' \in \mathcal{L}^l_j} WCRT^{l''}_j \tag{19}$$

Fig. 5 shows a diagram of the MTS^l_j calculation.

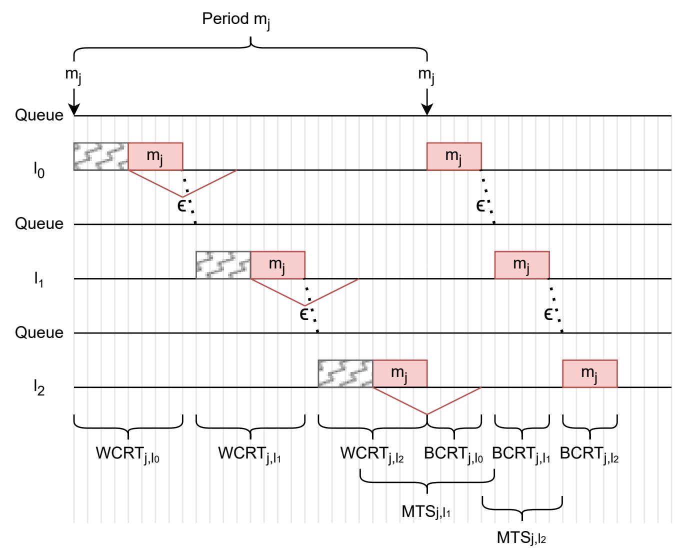


Fig. 5: Minimum temporal separation between two consecutive frames of the same stream.

Proof. First, since the time difference includes the same number of switches ($|\mathcal{L}^l_j| - 1$), we can exclude the ϵ factor of the switches from the calculation. On the other hand, $D_j - C_j \times \frac{\alpha^{-}_{P_j,l}}{\alpha^{+}_{P_j,l}} - \sum_{\forall l'' \in \mathcal{L}^l_j} WCRT^{l''}_j$ calculates the time between the worst-case transmission plus the replenishment of the credit on link l' and the end of the period, while $(|\mathcal{L}^l_j| - 1) \times C_j$ calculates the best-case reception time of a frame by the output queue of link l. By combining both values we obtain the minimum temporal distance between two consecutive frames of the same stream, i.e. MTS^l_j. ■

Note that the transmission of same-priority streams before the arrival time of the frame under analysis does not apply to the first link in the path of the frame under analysis, i.e., $\mathcal{L}_i(0)$. For the first link, we will use the pessimistic assumption from the previous analysis (Fig. 2) since we cannot guarantee that the transmitter end-station application will not attempt to send all same-priority frames simultaneously. Consequently, it is also unnecessary to consider the MTS^l_j of frames that share the same source as the frame under analysis. Due to the constraint deadline condition, all their previous frames should have already been received by the time the transmission of the frame under analysis begins.

Lemma 7.3: The minimum time the link l will be able to transmit same-priority traffic as the frame under analysis m_i (MTT_i^l) is:

$$MTT_i^l = \min\left(MRT_i^l - C_{P_i,l}^{max} \times \frac{\alpha_{P_i,l}^-}{\alpha_{P_i,l}^+}, \min_{\substack{\forall m_j \in sp(m_i) \\ \wedge \exists \mathcal{L}_j(x)=\mathcal{L}_i(y)=l \\ \wedge \mathcal{L}_j(0)\neq\mathcal{L}_i(0)}} \left(MTS_j^l \right) \right) \tag{20}$$

Proof. Firstly, in the absence of previous frames from streams of the same priority as the frame under analysis, the minimum time that link l can transmit traffic of the same priority as the frame under analysis m_i is MRT_i^l minus the time necessary to recover the minimum credit of l, denoted as $C_{P_i,l}^{max} \times \frac{\alpha_{P_i,l}^-}{\alpha_{P_i,l}^+}$. This is because, in the worst-case scenario, we assume that at the beginning of the MRT_i^l, the credit is at its minimum, thereby limiting the transmission capacity of same-priority frames during the MRT_i^l. When same-priority streams with different sources converge on the path of the frame under analysis, part of the MRT_i^l will, in the worst case, be allocated to the retransmission of preceding frames.

For a frame to interfere with the frame under analysis, it must arrive at least just before the frame under analysis. Furthermore, Lemma 7.2 demonstrates that there is a minimum time interval between a frame and its predecessor. Consequently, in the worst-case scenario, the frame with the minimum MTS_j^l will have arrived just before the frame under analysis, implying that its predecessor frame was transmitted through link l at least MTS_j^l time units earlier. If the minimum MTS_j^l is bigger than $MRT_i^l - C_{P_i,l}^{max} \times \frac{\alpha_{P_i,l}^-}{\alpha_{P_i,l}^+}$ then link l will be able to transmit same-priority traffic during the whole $MTS_j^l - C_{P_i,l}^{max} \times \frac{\alpha_{P_i,l}^-}{\alpha_{P_i,l}^+}$; otherwise, previous frames of stream m_j leaves only MTS_j^l for the transmission of frames that may interfere with the frame under analysis. ■

Lemma 7.4: In the absence of blocking and interference from other priorities, the transmission time of same-priority traffic MTT_i^l results in a reduction of SPI equivalent to its value, provided that it is either greater than or equal to 0 or less than or equal to SPI.

Proof. During MTT_i^l in the absence of blocking and interference from other priorities, at least $MTT_i^l \times \frac{\alpha_{P_i,l}^+}{\alpha_{P_i,l}^+ + \alpha_{P_i,l}^-}$ same-priority frames will be forwarded, leading to a SPI delay reduction of

$$X \times \frac{\alpha_{P_i,l}^+}{\alpha_{P_i,l}^+ + \alpha_{P_i,l}^-} \times \left(1 + \frac{\alpha_{P_i,l}^-}{\alpha_{P_i,l}^+}\right) = X \tag{21}$$

of the frame under analysis. Consequently, MTT_i^l corresponds to SPI component reduction. ■

Therefore, the maximum SPI, assuming no blocking or interference, is:

$$NewSPI_i^l = SPI_i^l - \min\left(\max\left(MTT_i^l, 0\right), SPI_i^l\right) \tag{22}$$

B. SPI with Blocking and Interference from other Classes

The new analysis is divided into two phases: phase 1 (P1) corresponds to the transmission of same-priority frames (MTT_i^l) occurring before the reception of the frame under analysis, while the second phase (P2) corresponds to the blocking and interference affecting the frame under analysis after it has arrived at the transmission queue.

Lemma 7.5: Any blocking or interference affecting link l during P1 has the same effect as if it had occurred during P2.

Proof. Blocking or interference affecting link l during P1 will reduce the MTT_i^l transmission of same-priority frames by a certain amount of time. This reduction increases the $NewSPI_i^l$ (Eq. (22)) by the same amount, up to a maximum of MTT_i^l as demonstrated in the proof of Lemma 7.4. ■

As demonstrated in [10], the maximum blocking and interference from non-ST traffic experienced by the transmission queue, provided there is pending traffic, is defined by Eq. (6). Additionally, while same-priority traffic continues to accumulate in the transmission queue and until the frame under analysis is transmitted, the queue will consistently contain traffic, thus keeping the maximum levels of non-ST blocking and interference as in Eq. (6) for the combined phases P1 and P2. In other words, non-ST blocking and interference remain the same regardless of the phase in which it occurs (P1 or P2) and are therefore independent of the improvement.

On the other hand, it is necessary to account for the STI during MTT_i^l. Thus, when computing the $WCRT_{i,c[k]}^{l,(x)}$ in Eq. (15) using the new SPI value (i.e., $NewSPI$ from Eq. (22), the MTT_i^l value should be added to calculate the STI. After obtaining $WCRT_{i,c[k]}^{l,(x)}$ iteratively, the MTT_i^l value would then be subtracted again, as MTT_i^l occur before the reception of the frame under analysis and, therefore, do not count for the $WCRT_{i,c[k]}^{l,(x)}$ calculation. This is equivalent to calculating the $WCRT_i^l$ as in the previous analysis (Eq. (3)) and subsequently subtracting MTT_i^l, i.e.:

$$newWCRT_i^l = WCRT_i^l - \min\left(\max\left(MTT_i^l, 0\right), SPI_i^l\right) \tag{23}$$

However, MTT_i^l depends on MTS_j^l which depends on $newWCRT_i^l$. Therefore, we will start by calculating $WCRT_i^l$ and MRT_i^l for all AVB frames. Next, we will calculate $newWCRT_i^{l,(0)}$ without considering MTS_j^l for all AVB frames. Finally, we iterate using the formula

$$\begin{aligned} &newWCRT_i^{l,(x)} = \\ &\quad WCRT_i^l - \\ &\quad \min\left(\max\left(MTT_i^l\left(newWCRT_i^{l,(x-1)}\right), 0\right), SPI_i^l\right) \end{aligned} \tag{24}$$

for all AVB frames until $newWCRT_i^{l,(x)} = newWCRT_i^{l,(x-1)}$ for all frames. Once this is achieved, we calculate $WCRT_i$ as in Eq. (4) by substituting $WCRT_i^l$ with $newWCRT_i^l$.

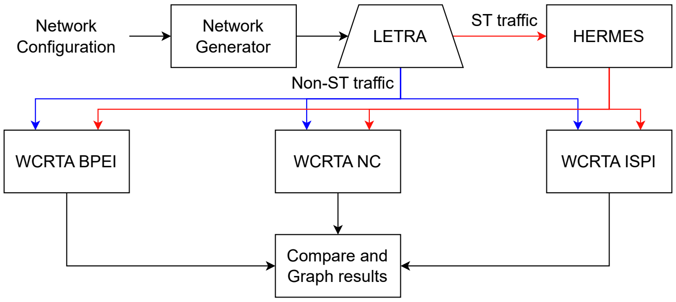


Fig. 6: LETRA configuration.

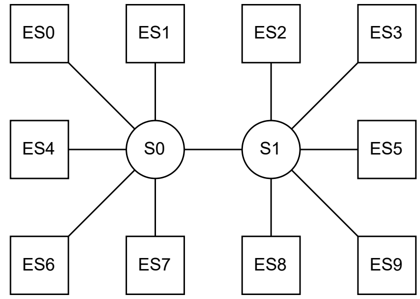


Fig. 7: Experimental network topology N1.

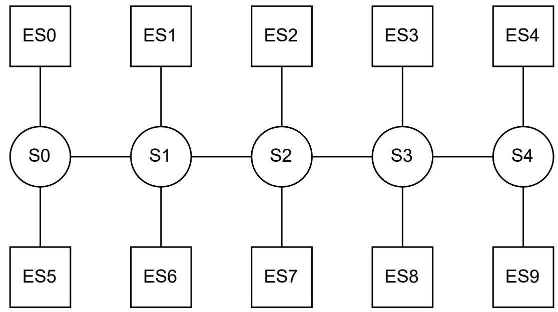


Fig. 8: Experimental network topology N2.

VIII. Experimental Setup

This study utilizes the LETRA Evaluation Toolset [17] to assess the proposed SPI enhancement. LETRA is an extensive suite of integrated tools designed for automated experiments, focusing on the scheduling and schedulability analysis of TSN networks. This section outlines the LETRA configuration used in this research, including specific modifications made for this study. The configuration is depicted in Fig. 6. The input for the evaluation toolset includes the network's configuration, encompassing its topology and traffic characteristics.

We examine two network topologies, illustrated in Figs. 7 and 8, which follow a line-star topology. This topology is suitable for our analysis, as the only missing element that could affect the results is the presence of loops. However, since some of the compared WCRTAs do not support circular dependencies, loops were excluded to ensure a fair comparison. Network N1 consists of a compact network with 2 switches, each connected to 5 end-stations, while Network N2 features a larger network with 5 switches, each connected to 2 end-stations. These topologies are part of the LETRA input.

To keep experiment durations manageable, the network bandwidth was set to 100 Mbps. This setting ensures that the maximum allowed 300 frames can consistently reach the target utilization on nearly every link. Frame lengths were chosen from the range [500, 1500] B. The minimum and maximum allowed periods were set at 10,000 μs and 30,000 μs, respectively.

LETRA begins with the Network Generator, which creates random traffic based on the provided topology and traffic characteristics. We enforced a traffic distribution of 5% ST and 95% AVB Class A and Class B. For the experiments, the BE class and AVB priorities lower than Class B were omitted due to limitations in the compared WCRTAs. We evaluated the performance of the WCRTAs across various network utilizations, ranging from 5% to 45%. We conducted 100 traffic generations for each utilization level, resulting in 900 experiments. Each experiment involved analyzing up to 300 frames, totaling nearly 270,000 frames.

In the next step, the generated traffic is mapped into the different TSN traffic classes (ST, AVB Class A, and AVB Class B) using the Mapping Tool (Fig. 6). The ST traffic is scheduled using an existing heuristic algorithm [14], although any other ST scheduling algorithm could be used. We chose a heuristic algorithm for its balance between speed and feasibility.

The AVB traffic and the ST schedule serve as inputs for each of the compared WCRTAs. The WCRTAs compared are: the WCRTA based on busy period and eligible interval (BPEI) from [3], the WCRTA based on Network Calculus (NC) from [4], and the WCRTA with improved SPI (ISPI) proposed in this paper, which extends the BPEI method with a new SPI calculation. All WCRTAs were configured with AVB Classes' credit slopes ($\alpha^+_{P_A,l}$, $\alpha^-_{P_A,l}$, $\alpha^+_{P_B,l}$, and $\alpha^-_{P_B,l}$) set to 0.5, equally dividing the available bandwidth between AVB Classes A and B, in line with the experimental setup. Additionally, a switch traversal factor ϵ of 0 was used to ensure a fair comparison, as some analyses do not consider this factor.

Finally, the WCRTAs are compared using two methods. First, the schedulability of each WCRTA is determined for each bandwidth utilization, calculated as the percentage of generated networks that meet their time requirements according to each WCRTA. Second, the pessimism ratio between each previous WCRTA and the one proposed in this paper is calculated for each bandwidth utilization. Specifically, for each AVB-generated frame m_i, the pessimism ratio of WCRTA X to WCRTA ISPI is computed as follows:

$$\textit{Pessimism Ratio}_i = \frac{WCRT\ X_i}{WCRT\ ISPI_i} \quad (25)$$

This value demonstrates the extent to which our proposed SPI improvement (WCRTA ISPI) reduces pessimism compared to the analyses named X, specifically BPEI and NC. A pessimism ratio below 1 indicates an increase in pessimism in the SPI calculation, while a ratio above 1 indicates the degree to which the proposed SPI improvement reduces pessimism.

Given the random traffic generation and the large number of experiments conducted, we can confidently state that these experiments do not favor any particular WCRTA. Consequently, the results accurately reflect the performance differences among the compared WCRTAs under the specified topologies and traffic characteristics. Although the improvement over WCRTA BPEI is analytically validated, ensuring the improvement over WCRTA NC in all TSN topologies and traffic configurations is more challenging due to the inherent differences between the analyses.

IX. Results and Discussion

This section presents the results of the experiments described in Section VIII. We begin by presenting the schedulability results of the three WCRTAs (BPEI, NC, and ISPI) across the two network topologies for the various utilization percentages. Following this, we analyze the additional pessimism observed in WCRTA BPEI and NC compared to WCRTA ISPI, as this higher pessimism contributes to their lower schedulability.

A. Schedulability Results

Figs. 9 and 10 illustrate the schedulability percentages achieved by each evaluated WCRTA across different bandwidth utilizations for networks N1 and N2, respectively.

The results clearly demonstrate that the WCRTA incorporating the improved SPI calculation introduced in this study, consistently outperforms the previous WCRTAs in terms of schedulability. Specifically, the proposed WCRTA ISPI achieves up to 90% higher schedulability compared to WCRTA BPEI and up to 40% higher schedulability at certain utilization levels when compared to WCRTA NC. It is also important to highlight that, despite the general decrease in network schedulability with an increase in network hops (Fig. 10), the proposed WCRTA continues to exhibit superior performance over the previous solutions.

B. Additional Pessimism

Fig. 11 illustrates the additional pessimism observed in the WCRTA BPEI and NC compared to the WCRTA ISPI across all analyzed networks and utilization. The figures show box plots formed using the pessimism ratios of all the frames analyzed for each schedulable utilization percentage. It is important to note that the WCRT values are only valid for schedulable networks, as the studied WCRTAs are based on the constrained deadline condition. Therefore, if any frame misses its deadline, the WCRT value becomes unreliable. The x-axis represents different utilization percentages, while the y-axis indicates the pessimism ratio. A horizontal dashed line is included at a pessimism ratio of 1, serving as a reference point. Values above 1 signify greater pessimism in previously proposed WCRTAs compared to the proposed solution, whereas values below 1 indicate that the proposed solution introduces more pessimism relative to the previously

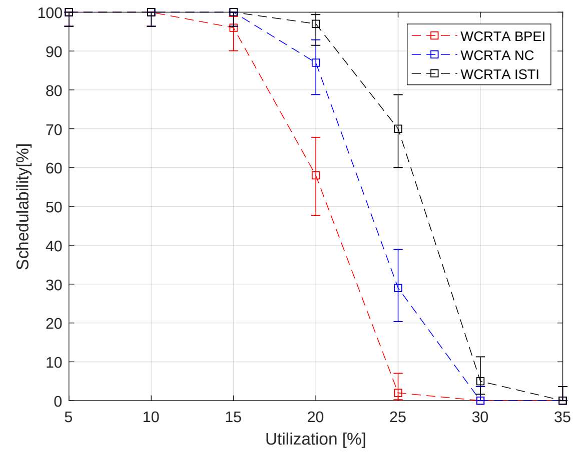


Fig. 9: Schedulability of the WCRTAs in network N1.

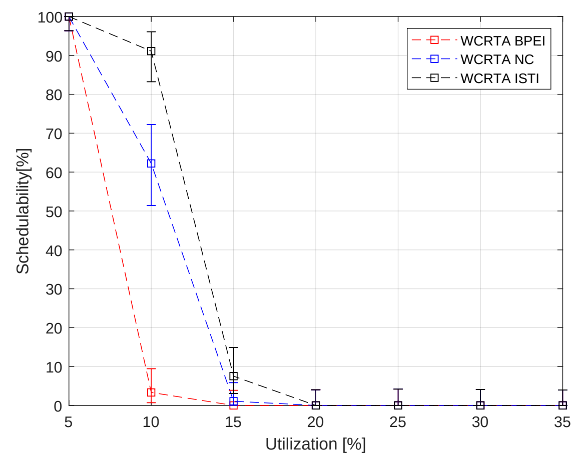


Fig. 10: Schedulability of the WCRTAs in network N2.

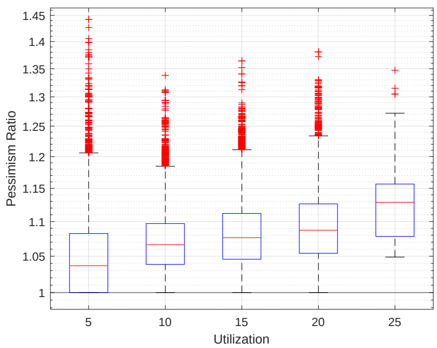
compared WCRTA (BPEI or NC). The figure clearly demonstrates that, in all cases, the BPEI and NC WCRTAs exhibit greater pessimism than the WCRTA ISPI, which accounts for the higher schedulability of the latter.

X. Conclusion

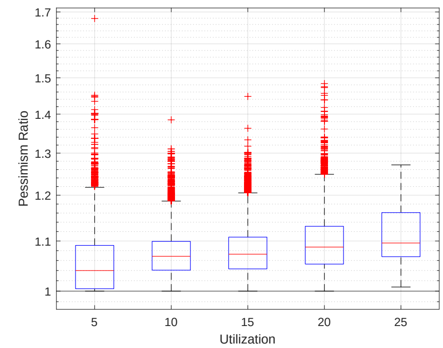
Reducing pessimism in current Audio-Video Bridging (AVB) Worst-Case Response Time Analysis (WCRTA) is crucial for enhancing the practicality of Time-Sensitive Networks (TSN) and, consequently, its adoption by the industry. This paper addresses one of the primary sources of pessimism in existing analyses: the Same-priority Interference (SPI). We demonstrate that in TSN, AVB frames do not need to be interfered with by all same-priority frames at each switch, but rather by a smaller subset of them. This significant reduction in pessimism leads to an increase in the schedulability of the analysis when using the improved SPI calculation, compared to those that do not incorporate this improvement.

Acknowledgment

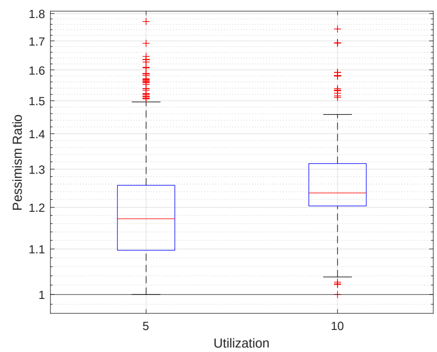
This research is supported by the Swedish Knowledge Foundation (KKS) under the SEINE and MARC projects, and



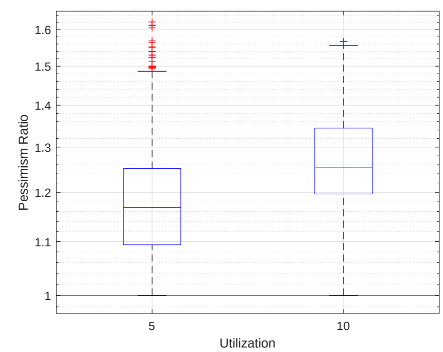
(a) AVB Class A traffic of WCRTA BPEI on networks N1



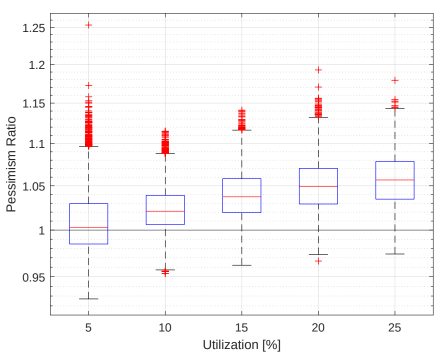
(b) AVB Class B traffic of WCRTA BPEI on networks N1



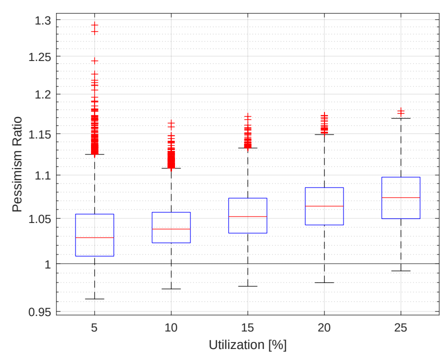
(c) AVB Class A traffic of WCRTA BPEI on networks N2



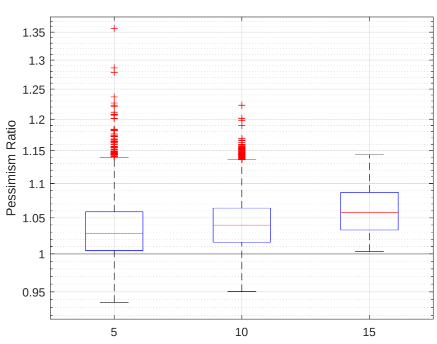
(d) AVB Class B traffic of WCRTA BPEI on networks N2



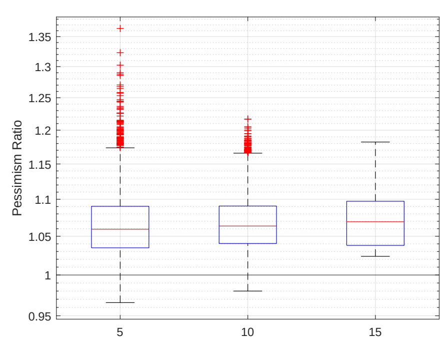
(e) AVB Class A traffic of WCRTA NC on networks N1



(f) AVB Class B traffic of WCRTA NC on networks N1



(g) AVB Class A traffic of WCRTA NC on networks N2



(h) AVB Class B traffic of WCRTA NC on networks N2

Fig. 11: Additional pessimism observed in the WCRTA BPEI and NC compared to the WCRTA ISPI across all analyzed networks and utilizations.

the Swedish Governmental Agency for Innovation Systems (VINNOVA) under the FLEXATION and iSecure projects. Julian Proenza was supported by Grant pid2021-124348ob-i00, funded by MCIN/AEI/ 10.13039/501100011033 / ERDF, EU.

REFERENCES

[1] *IEEE 802.1Q-2022: IEEE Standard for Local and Metropolitan Area Networks–Bridges and Bridged Networks*, Std., 2022. [Online]. Available: https://standards.ieee.org/ieee/802.1Q/10323/

[2] Bujosa, Daniel and Proenza, Julián and Papadopoulos, Alessandro V and Nolte, Thomas and Ashjaei, Mohammad, "An Improved Worst-Case Response Time Analysis for AVB Traffic in Time-Sensitive Networks," in *Proceedings-Real-Time Systems Symposium*, 2024.

[3] L. Lo Bello, M. Ashjaei, G. Patti, and M. Behnam, "Schedulability analysis of Time-Sensitive Networks with scheduled traffic and preemption support," *Journal of Parallel and Distributed Computing*, vol. 144, pp. 153–171, 2020.

[4] L. Zhao, P. Pop, Z. Zheng, H. Daigmorte, and M. Boyer, "Latency analysis of multiple classes of AVB traffic in TSN with standard credit behavior using network calculus," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 10, pp. 10 291–10 302, 2020.

[5] J. Diemer, J. Rox, and R. Ernst, "Modeling of Ethernet AVB networks for worst-case timing analysis," *IFAC Proceedings Volumes*, vol. 45, no. 2, pp. 848–853, 2012.

[6] J. Diemer, D. Thiele, and R. Ernst, "Formal worst-case timing analysis of Ethernet topologies with strict-priority and AVB switching," in *7th IEEE International Symposium on Industrial Embedded Systems (SIES'12)*. IEEE, 2012, pp. 1–10.

[7] U. D. Bordoloi, A. Aminifar, P. Eles, and Z. Peng, "Schedulability analysis of Ethernet AVB switches," in *2014 IEEE 20th International Conference on Embedded and Real-Time Computing Systems and Applications*. IEEE, 2014, pp. 1–10.

[8] Cao, Jingyue and Cuijpers, Pieter JL and Bril, Reinder J and Lukkien, Johan J, "Tight worst-case response-time analysis for ethernet AVB using eligible intervals," in *2016 IEEE World Conference on Factory Communication Systems (WFCS)*. IEEE, 2016, pp. 1–8.

[9] ——, "Independent yet tight WCRT analysis for individual priority classes in ethernet AVB," in *Proceedings of the 24th International Conference on Real-Time Networks and Systems*, 2016, pp. 55–64.

[10] ——, "Independent WCRT analysis for individual priority classes in Ethernet AVB," *Real-Time Systems*, vol. 54, pp. 861–911, 2018.

[11] Maxim, Dorin and Song, Ye-Qiong, "Delay analysis of AVB traffic in time-sensitive networks (TSN)," in *Proceedings of the 25th International Conference on Real-Time Networks and Systems*, 2017, pp. 18–27.

[12] L. Zhao, P. Pop, Z. Zheng, and Q. Li, "Timing analysis of AVB traffic in TSN networks using network calculus," in *2018 IEEE Real-Time and Embedded Technology and Applications Symposium (RTAS)*. IEEE, 2018, pp. 25–36.

[13] M. Seliem, A. Zahran, and D. Pesch, "Delay analysis of TSN based industrial networks with preemptive traffic using network calculus," in *2023 IFIP Networking Conference (IFIP Networking)*. IEEE, 2023, pp. 1–9.

[14] Bujosa, Daniel and Ashjaei, Mohammad and Papadopoulos, Alessandro V and Nolte, Thomas and Proenza, Julián, "HERMES: Heuristic multi-queue scheduler for TSN time-triggered traffic with zero reception jitter capabilities," in *Proceedings of the 30th International Conference on Real-Time Networks and Systems*, 2022, pp. 70–80.

[15] R. I. Davis, S. Kollmann, V. Pollex, and F. Slomka, "Controller Area Network (CAN) Schedulability Analysis with FIFO Queues," in *2011 23rd Euromicro Conference on Real-Time Systems*. IEEE, 2011, pp. 45–56.

[16] J. Maki-Turja and M. Nolin, "Fast and tight response-times for tasks with offsets," in *17th Euromicro Conference on Real-Time Systems (ECRTS'05)*, 2005.

[17] Mateu, Daniel Bujosa and Ashjaei, Mohammad and Papadopoulos, Alessandro V and Proenza, Julian and Nolte, Thomas, "LETRA: Mapping legacy ethernet-based traffic into TSN traffic classes," in *2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. IEEE, 2021, pp. 1–8.