



Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <http://oatao.univ-toulouse.fr/>
Eprints ID: 11568

To cite this document: Dang, Dinh Khanh and Mifdaoui, Ahlem *Timing Analysis of TDMA-based Networks using Network Calculus and Integer Linear Programming*. (2014) In: IEEE 22nd International Symposium on Modeling Analysis and Simulation of Computer and Telecommunication Systems (MASCOTS), 9 September 2014 - 11 September 2014 (Paris, France).

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@inp-toulouse.fr

Timing Analysis of TDMA-based Networks using Network Calculus and Integer Linear Programming

Dinh-Khanh Dang
University of Toulouse-ISAE
dinh-khanh.dang@isae.fr

Ahlem Mifdaoui
University of Toulouse-ISAE
ahlem.mifdaoui@isae.fr

Abstract—For distributed safety-critical systems, such as avionics and automotive, shared networks represent a bottleneck for timing predictability, a key issue to fulfill certification requirements. To control interferences on such shared resources and guarantee bounded delays, the Time Division Multiple Access (TDMA) protocol is considered as one of the most interesting arbitration protocols due to its deterministic timing behavior and fault-tolerance features. This paper addresses the problem of computing the worst-case end-to-end delay bounds for traffic flows sharing a TDMA-based network using Network Calculus. First, we extend classic timing analysis to integrate the impact of non-preemptive message transmission and various service policies in end-systems, e.g., First In First Out (FIFO), Fixed Priority (FP) and Weighted Round Robin (WRR). Afterwards, the proposed models are refined using Integer Linear Programming (ILP) to obtain tighter end-to-end delay bounds. Finally, this general analysis is illustrated and validated in the case of a TDMA-based Ethernet network for I/O avionics applications. Results show the efficiency of the proposed models to provide stronger guarantees on system schedulability, compared to classic models.

I. INTRODUCTION

For safety-critical applications, such as avionic and automotive, guarantees on worst-case behavior are key issues to fulfill certification requirements. The main difficulty to conduct such an analysis arises from the increasing complexity of new generation safety-critical systems. These new systems are generally based on shared resources, such as networks for distributed end-systems. These shared networks represent the bottleneck for performance and timing predictability. To control interferences and guarantee bounded access delays, various arbitration mechanisms can be integrated.

Among existing arbitration protocols, Time Division Multiple Access (TDMA) has been successfully used in various embedded networks, such as FlexRay [3] for automotive and more recently Time Triggered Ethernet [6] for avionics. TDMA consists of partitioning the network access over time, and assigning it to only one end-system during any particular time slot. The use of TDMA to control medium-access has many advantages due to its deterministic timing behavior and its contention-free and fault-tolerance features. However, TDMA may lead at the same time to increasing communication latencies, thus requiring real-time constraints be verified. To deal with the worst-case performance analysis of such systems, an appropriate timing analysis to provide worst-case end-to-end delay bounds has to be considered.

Many challenges arise from conducting such an analysis. First, safety-critical systems are based on non-preemptive

communication where a message transmission cannot straddle two slots, and consequently if the remaining time during a slot is not enough for a complete transmission, then the message has to wait for the next slot. Second, the communication networks in such systems are used to transmit traffic flows with different temporal constraints. This implies guaranteeing a delay bound for each data-type and proving the temporal isolation between flows. Third, the end-systems sharing the TDMA-based network implement service policies, such as First In First Out (FIFO), Fixed Priority (FP) and Weighted Round Robin (WRR) policies. The impact of such policies on the analysis of worst-case end-to-end delays needs to be integrated.

Among analytical methods to conduct worst-case performance analysis of distributed networks, Network Calculus [8] is well adapted to controlled traffic sources and provides upper bounds on delays for traffic flows. It has been used in many application fields, such as wireless sensor networks [5] [7], avionics [9] and space networks [4]. In the area of timing analysis for TDMA-based systems, some interesting work based on Network Calculus has been proposed [5] [7] [10]. These approaches are mainly based on fluid flow models which may result in optimistic end-to-end delay bounds, i.e., less than the delays that may actually occur. This fact may provide guarantees for messages that will actually miss their deadlines in the worst-case when considering packet flow models.

The contributions of this work are: (i) Efficient timing analysis using Network Calculus for a packet flow model under FIFO, FP and WRR policies sharing a TDMA-based network. To our knowledge, WRR modeling under these conditions has not been treated in the literature; (ii) Refining the proposed models for the different service policies using Integer Linear Programming (ILP) to compute tighter end-to-end delay bounds; (iii) The validation of such timing analyses in the case of a TDMA-based Ethernet network for I/O avionics applications. The efficiency of the proposed models to provide stronger guarantees on system schedulability is proved, with reference to classic models.

In the next section, we describe the basic concepts of Network Calculus framework, and we review the most relevant work in the area of timing analysis using Network Calculus of TDMA-based networks. Afterwards, the timing analysis of TDMA-based networks under various service policies is tackled as follows. First, the system modeling and schedulability analysis methodology are presented in Section III. Then, extended service curves based on packet flow models under various service policies are detailed in Section IV, and refined

using ILP in Section V. Finally, the validation of such timing analysis is illustrated within a realistic avionic application.

II. BACKGROUND AND RELATED WORK

A. Network Calculus Framework

The timing analysis of TDMA-based networks detailed in this paper is based on *Network Calculus* theory [8], providing upper bounds on delays and backlogs. Delay bounds depend on the traffic arrival described by the so called *arrival curve* α , and on the availability of the traversed node described by the so called *minimum simple or strict service curve* β . The definitions of these curves are explained as following.

Definition 1. (*Arrival Curve*) a function α is an arrival curve for a data flow with an input cumulative function R , such that $R(t)$ is the number of bits received until time t , iff:

$$\forall t, s \geq 0, s \leq t, R(t) - R(s) \leq \alpha(t - s)$$

Definition 2. (*Simple service curve*) The function β is the minimum simple service curve for a data flow with an input cumulative function R and output cumulative function R^* iff:

$$R^* \geq R \otimes \beta$$

where $(f \otimes g)(t) = \inf_{0 \leq s \leq t} \{f(t - s) + g(s)\}$

Definition 3. (*Strict service curve*) The function β is the minimum strict service curve for a data flow with an input cumulative function R and output cumulative function R^* , if for any backlogged period $]s, t]^1$, $R^*(t) - R^*(s) \geq \beta(t - s)$.

The delay bound D is the maximum horizontal distance between α and β called $h(\alpha, \beta)$; whereas the backlog bound B is the maximum vertical distance called $v(\alpha, \beta)$. Finally, to compute delay bounds of individual traffic flows, we need the residual service curve according to the following theorem.

Theorem 1. (*Residual service curve - Blind Multiplex*) [1] let f_1 and f_2 be two flows crossing a server that offers a strict service curve β such that f_1 is α_1 -constrained, then the residual service curve offered to f_2 is:

$$\beta_2 = (\beta - \alpha_1)_\uparrow$$

where $f_\uparrow(t) = \max\{0, \sup_{0 \leq s \leq t} f(s)\}$

B. Timing Analysis of TDMA-based Networks

A timing analysis of TDMA-based systems using Network Calculus aims to provide a method to compute end-to end delay bounds of transmitted messages. These obtained values are then compared to message deadlines to verify system schedulability.

In [5] [7], the authors applied Network Calculus to provide real-time guarantees for Wireless Sensor Networks (WSNs). The former work proposed an optimization approach to design a TDMA arbiter for generic sink-tree WSNs; whereas the latter focused on performance analysis of such networks considering the sink mobility. Another interesting work in this area was proposed by Wandeler et al. [10] where the authors proposed an approach to find the optimal cycle length as well as

the minimum required bandwidth of TDMA resource under FIFO and FP service policies in end-systems. These different approaches are based on a fluid flow model which may lead to optimistic end-to-end delays when considering a packet flow model with non-preemptive message transmission.

In this paper, we provide extended timing analysis of such systems to integrate the non-preemption of message transmission under various service policies such as FIFO, FP and WRR. Further, we refine the proposed models using Integer Linear Programming (ILP), using the same idea introduced in [2] for tandem networks under FIFO multiplexing, to compute tighter end-to-end delay bounds. Finally, we discuss the possible impact of such timing analysis in the case of a TDMA-based Ethernet network for I/O avionics applications. The efficiency of our proposed models to provide stronger guarantees on system schedulability compared to classic models is validated.

III. TIMING ANALYSIS METHODOLOGY

To conduct timing analysis of TDMA-based Networks using Network Calculus, we first define the system scenario and assumptions. Then, we present the considered schedulability test and detail models of traffic flows, end-systems and the TDMA arbitration protocol.

A. System Scenario and Assumptions

The considered end-systems generate messages independently and transmit their generated traffic flows on the shared medium under various service policies: FIFO, FP or WRR.

The inter-communication between these end-systems is based on a TDMA network and follows a static TDMA schedule. A TDMA schedule is defined as a sequence of time slots repeated each cycle with a fixed duration, called c . During each cycle, each end-system can only transmit during a predetermined time interval, called TDMA time slot s . We consider the general case where the time slots associated to end-systems have not necessarily equal durations. The message transmission on this TDMA-based network is non-preemptive. Consequently, if the remaining time during a slot is insufficient for a complete message transmission, then the message has to wait for the next slot. Hence, the cycle duration is as follows:

$$c = \sum_{k=1}^M s^k + t_{sync} \quad (1)$$

where M is the number of end-systems, and t_{sync} the duration of the synchronization phase.

B. Sufficient Schedulability Test

Using Network Calculus, an upper bound on end-to-end delay of each transmitted message m , D_m^{eed} , is computed and then compared to its respective temporal deadline, DI_m . Hence, this schedulability test results in a sufficient but not necessary condition due to the pessimism introduced by the upper bounds. Nevertheless, we can still infer the traffic schedulability as follows:

$\forall m \in \text{messages},$

$$D_m^{eed} \leq DI_m \implies \text{The messages set is schedulable}$$

¹ $]s, t]$ is called backlogged period if $R(\tau) - R^*(\tau) > 0, \forall \tau \in]s, t]$

The end-to-end delay of each transmitted message m consists mainly of three parts:

- arbitration delay that corresponds to the maximum waiting time since the arrival instant of the message until it starts being transmitted on the network medium. This delay is due to interference caused by other messages from the same end-system, and the waiting time due to TDMA arbitration protocol;
- transmission time that corresponds to the communication time of the message on the medium; it depends on the message size and the medium transmission capacity;
- propagation delay needed to propagate a signal from a source to its final destinations. In our case, this delay is considered insignificant.

C. System Modeling

To compute upper bounds on end-to-end delays of transmitted messages using Network Calculus, we need to model each message flow to compute its maximum arrival curve, and the behavior of end-systems and the TDMA arbitration protocol to compute the minimum service curve.

The characteristics of an aggregate traffic flow f_i generated by an end-system are defined as $(n_i, T_i, D_i, L_i, e_i)$ for the number of message flows, the message period (and minimum inter-arrival time for sporadic flow), deadline (equal to T_i unless otherwise explicitly specified), frame size integrating the protocol overhead and delivery time (i.e., $e_i = L_i/B$ where B is the medium transmission capacity), respectively.

The arrival curve of traffic flow i , based on a packetized model is as follows.

$$\alpha_i(t) = n_i L_i \lceil \frac{t}{T_i} \rceil \quad (2)$$

For end-systems and TDMA protocol modeling, first we present existent classic models based on a fluid flow model for FIFO and FP policies. However, to our best knowledge, no model for WRR policy combined with TDMA protocol currently exists. These models will be then extended in the next sections to integrate the impact of non-preemptive message transmission.

The classic service curve for a fluid flow model when FIFO policy is implemented in the end-system has the following analytical expression:

$$\beta_{c,s}(t) = B \max(\lfloor \frac{t}{c} \rfloor s, t - \lceil \frac{t}{c} \rceil (c-s)), \forall t \geq 0 \quad (3)$$

The main idea is based on the fact that an end-system with a time slot s may not have access to the shared network during at maximum $c-s$. After this maximum duration, the end-system has exclusive access to the medium during its time slot s to transmit with the medium transmission capacity, B . When considering FP policy, each traffic flow will be transmitted before all lower priority flows and after all higher priority flows. Consider N traffic flows f_1, \dots, f_N where f_i has higher priority than f_j if $i < j$. The residual service curve offered to

traffic flow f_i using Theorem 1 has the following analytical expression:

$$\beta_i(t) = (\beta_{c,s}(t) - \sum_{1 \leq j \leq i-1} \alpha_j(t))_{\uparrow} \quad (4)$$

D. Computing Upper Bounds on End-to-End Delays

Upper bounds on end-to-end delays can easily be computed as the maximum horizontal distance between the maximum arrival curve of the traffic and the minimum service curves guaranteed by each end-system when implementing FIFO, FP or WRR policies.

Consider an end-system generating N traffic flows and each flow f_i characterized with the maximum arrival curve $\alpha_i(t)$, the upper bound end-to-end delay D_i^{eed} associated to flow f_i depends on the implemented service policy in the end-system:

- with FIFO policy, D_i^{eed} is the upper bound on end-to-end delay guaranteed to the aggregate traffic flow generated by the same end-system when considering the minimum service curve $\beta(t)$. Hence,

$$\forall i D_i^{eed} = h(\alpha, \beta) \text{ where } \alpha(t) = \sum_{i=1}^N \alpha_i(t) \quad (5)$$

- with FP and WRR policies, D_i^{eed} is the maximum horizontal distance between the maximum arrival curve $\alpha_i(t)$ and the guaranteed residual service curve β_i . Hence,

$$\forall i D_i^{eed} = h(\alpha_i, \bar{\beta}_i) \quad (6)$$

IV. PACKET SERVICE CURVES USING NETWORK CALCULUS

In this section, the extended service curves based on a packet flow model are detailed and proved under various service policies, i.e., FIFO, FP and WRR.

A. FIFO Policy

The maximum waiting time to access the medium and the lower bound of offered TDMA time slot have to be adjusted in case of non-preemptive message transmission, and then used to extend the service curve guarantee.

Consider an end-system with a TDMA time slot s during each TDMA cycle c . This end-system generates N traffic flows with associated maximum and minimum delivery times e_{max} and e_{min} , respectively. Based on the worst-case scenario illustrated in Figure 1, the corresponding parameters are as following.

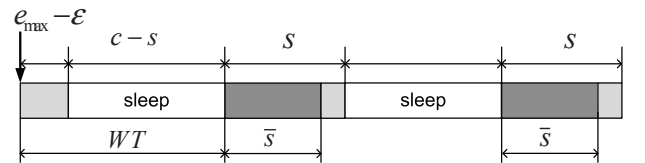


Figure 1. Worst case scenario with FIFO policy

- The maximum waiting time WT occurs when the first message of a backlogged period has a maximum delivery time

e_{max} and arrives just at the instant when the remaining time during the current slot is slightly less than the required delivery time, i.e., $e_{max} - \epsilon$ ($0 < \epsilon \ll 1$). In this case, the message has to be delayed until the next slot and the *maximum waiting time* is then $WT = e_{max} + c - s$ instead of $c - s$ in the classic model.

• The *lower bound of offered TDMA time slot* \bar{s} depends on message flows characteristics. If generated messages are homogeneous, i.e., all the messages have the same delivery time e , then the end-system can send at maximum $\lfloor \frac{s}{e} \rfloor$ messages during its respective time slot s . Else, the idle time is always less than e_{max} and the *lower bound of offered TDMA time slot* is always greater than e_{min} . Hence, \bar{s} is as follows:

$$\bar{s} = \begin{cases} \lfloor \frac{s}{e} \rfloor e & \text{if } e_{max} = e_{min} = e \\ \max\{s - e_{max}, e_{min}\} & \text{Otherwise} \end{cases} \quad (7)$$

The *maximum waiting time* WT may occur only at the beginning of the backlogged period. Then, during the next cycles the node may wait at most $c - \bar{s}$ to access to the network and transmit messages. Hence, the strict service curve guaranteed in this case is $\bar{\beta}(t) = \beta_{c, \bar{s}}(t - (WT - (c - \bar{s})))$ which represents the curve shifted to the right with the positive duration $WT - (c - \bar{s})$ of $\beta_{c, \bar{s}}(t)$ defined in Eq. 3.

This is explicitly defined in the following theorem. The corresponding proof is detailed in Appendix VIII-A.

Theorem 2. Consider an end-system having a lower bound of offered time slot \bar{s} , generating N traffic flows and implementing a FIFO policy. The offered strict service curve when considering non-preemptive message transmission is:

$$\bar{\beta}(t) = \beta_{c, \bar{s}}(t - WT + (c - \bar{s})), \forall t \geq 0$$

where

$$WT = e_{max} + c - s,$$

and

$$\bar{s} = \begin{cases} \lfloor \frac{s}{e} \rfloor e & \text{if } e_{max} = e_{min} = e \\ \max\{s - e_{max}, e_{min}\} & \text{Otherwise} \end{cases}$$

The extended service curve for a packet flow model is illustrated in Figure 2. Notice the extended service curve is lower than the classic one, described in Eq. 3. This difference leads to greater delay bounds, mainly due to the non-preemptive message transmission.

B. FP Policy

As for the FIFO policy case, the *maximum waiting time* to access to the medium and the *lower bound of offered TDMA time slot* for transmission have to be adjusted in case of non-preemptive message transmission. First, we compute the offered service curve for aggregate traffic flow $f_{<i}$ including traffic flows f_1, \dots, f_i with priorities higher or equal to i . Then, we deduce the individual service curve for each traffic flow f_i applying Theorem 1. Based on the worst-case scenario illustrated in Figure 3, the corresponding parameters are as follows:

• The *maximum waiting time* $WT_{<i}$ occurs when the first message of the aggregate message flow $f_{<i}$ arriving at the

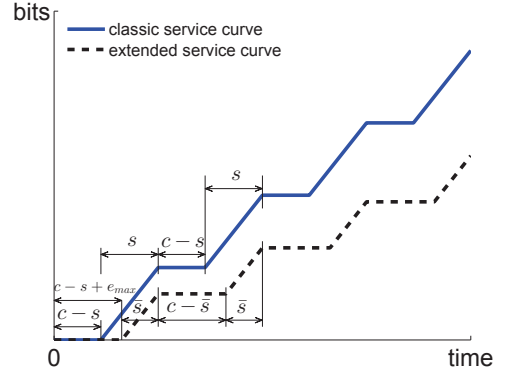


Figure 2. Classic vs extended service curves with FIFO policy

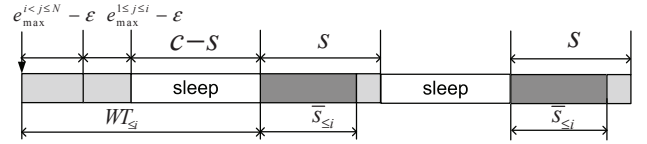


Figure 3. Worst case scenario with FP policy

beginning of a backlogged period and having a maximum delivery time $e_{max}^{1 \leq j \leq i}$ is blocked during the transmission time of a message with lower priority and maximum delivery time. This blocking time corresponds to $e_{max}^{j < i \leq N} - \epsilon$ ($0 < \epsilon \ll 1$). Furthermore, after that blocking time, the remaining time during the current slot is slightly less than the required delivery time of the considered message and is equal to $e_{max}^{1 \leq j \leq i} - \epsilon$ ($0 < \epsilon \ll 1$). In this case, this message has to wait until the next slot to be transmitted. The *maximum waiting time* is then bounded by $e_{max}^{j < i \leq N} + e_{max}^{1 \leq j \leq i} + c - s$. However, if TDMA time slot $s < e_{max}^{j < i \leq N} + e_{max}^{1 \leq j \leq i}$, then the *maximum waiting time* is reduced to one cycle $c (= s + c - s)$. Hence, for aggregate flow $f_{<i}$, the *maximum waiting time* is as follows:

$$WT_{<i} = \min(e_{max}^{j < i \leq N} + e_{max}^{1 \leq j \leq i} + c - s, c) \quad (8)$$

• The *lower bound of offered TDMA time slot* associated to aggregate message flow $f_{<i}$, called $\bar{s}_{<i}$, can be deduced from Eq. 7 by considering only aggregate flow $f_{<i}$ instead of all the generated flows. Hence, $\bar{s}_{<i}$ associated to aggregate flow $f_{<i}$ is as follows:

$$\bar{s}_{<i} = \begin{cases} \lfloor \frac{s}{e} \rfloor e & \text{if } e_{max}^{1 \leq j \leq i} = e_{min}^{1 \leq j \leq i} = e \\ \max(e_{min}^{1 \leq j \leq i}, s - e_{max}^{1 \leq j \leq i}) & \text{Otherwise} \end{cases} \quad (9)$$

The extended service curve with FP policy is defined in the following theorem and the mathematical proof is detailed in Appendix VIII-A.

Theorem 3. Consider aggregate traffic flow $f_{<i}$ having a lower bound of offered TDMA time slot $\bar{s}_{<i}$, transmitted by an end-system implementing FP policy. The strict service curve guaranteed to $f_{<i}$ when considering non-preemption feature is:

$$\bar{\beta}_{<i}(t) = \beta_{c, \bar{s}_{<i}}(t - WT_{<i} + (c - \bar{s}_{<i})), \forall t \geq 0 \quad (10)$$

where

$$WT_{\leq i} = \min(e_{\max}^{i < j \leq N} + e_{\max}^{1 \leq j \leq i} + c - s, c),$$

and

$$\bar{s}_{\leq i} = \begin{cases} \lfloor \frac{s}{e} \rfloor e & \text{if } e_{\max}^{1 \leq j \leq i} = e_{\min}^{1 \leq j \leq i} = e \\ \max(e_{\min}^{1 \leq j \leq i}, s - e_{\max}^{1 \leq j \leq i}) & \text{Otherwise} \end{cases}$$

Hence, using Theorems 1 and 3, the residual service curve offered to message flow f_i is as follows:

$$\bar{\beta}_i(t) = (\beta_{c, \bar{s}_{\leq i}}(t - WT_{\leq i} + (c - \bar{s}_{\leq i})) - \sum_{j=1}^{i-1} \alpha_j(t))_{\uparrow} \quad (11)$$

This extended service curve is illustrated in Figure 4. Extended service is lower than the classic service which leads to greater delay bounds due to the non-preemptive message transmission.

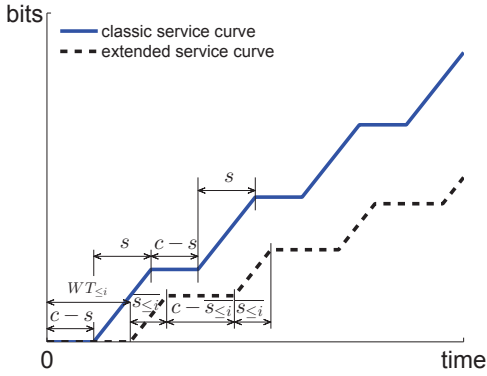


Figure 4. Classic vs extended service curves with FP policy

C. WRR Policy

As a first step, we detail a service curve model for WRR policy combined with TDMA protocol for a fluid flow model, which to our knowledge has not been investigated in literature. Then, we extend this model to the packet flow model.

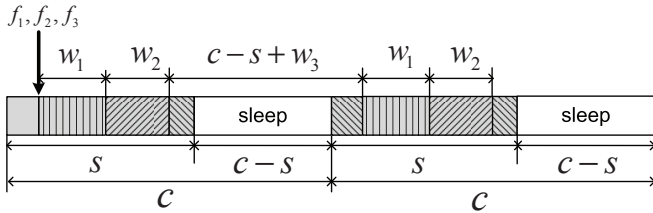


Figure 5. Access-time distribution with preemptive WRR over TDMA

As illustrated in Figure 5, for each flow f_i in an end-system, we consider a respective access time to the medium, called w_i during each WRR round, where $\sum_{i=1}^N w_i = s$. During a backlogged period of a flow f_i , the *maximum waiting time* occurs when it is blocked by all other flows having their backlogged period simultaneously. Hence, the *maximum waiting time* is bounded by $\sum_{j=1, j \neq i}^N w_j + c - s = c - w_i$. Then, flow f_i

receives the right to send during its respective access time w_i . This behavior is repeated according to weighted round-robin scheduling and the analytical expression is given in Theorem 4.

Theorem 4. The strict service curve for flow f_i transmitted by a node implementing a preemptive WRR combined with TDMA system, with a respective access time to the medium w_i such that $\sum_{i=1}^N w_i = s$ is

$$\beta_i(t) = \beta_{c, w_i}(t) \quad (12)$$

Proof: During any duration c , flow f_i can access and transmit messages during at least w_i . Hence, the minimum service curve of f_i is given by $\beta_i(t) = \beta_{c, w_i}(t)$. ■

Then, we extend this first model to integrate the non-preemption. This condition impacts the *maximum waiting time* and the respective access time to the medium w_i for each flow f_i . As shown in Figure 6,

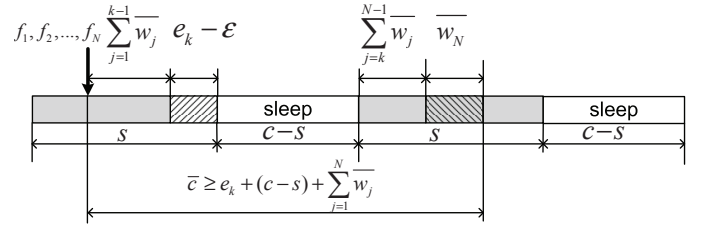


Figure 6. Worst case scenario with WRR policy

- The respective access time to the medium of flow f_i corresponds to the amount of time to transmit completely $\lfloor \frac{w_i}{e_i} \rfloor$ messages within a slot duration that is equal to $\bar{w}_i = \lfloor \frac{w_i}{e_i} \rfloor * e_i$.
- The *maximum waiting time* occurs when flow f_i starts its backlogged period and it is blocked by all the other backlogged flows during the slot. The *maximum waiting time* also has to integrate the remaining idle time between two slots due to flow f_k as shown in Figure 6, which is slightly less than e_{\max} . Consequently, the flow has to be delayed to the next slot. Hence, the *maximum waiting time* is bounded by $e_{\max} + c - s + \sum_{j=1, j \neq i}^N \bar{w}_j$. It is worth noting that during any period $\bar{c} = e_{\max} + c - s + \sum_{1 \leq j \leq N} \bar{w}_j$, the backlogged flow f_i can transmit at least $\lfloor \frac{w_i}{e_i} \rfloor$ messages.

The analytical expression of this offered service curve in this case is given in Theorem 5 and the proof is detailed in Appendix VIII-B.

Theorem 5. The strict service curve for f_i , transmitted by a node implementing a non-preemptive WRR combined with TDMA system is

$$\bar{\beta}_i(t) = \beta_{\bar{c}, \bar{w}_i}(t) \quad (13)$$

where $\bar{w}_i = \lfloor \frac{w_i}{e_i} \rfloor * e_i$, and $\bar{c} = e_{\max} + (c - s) + \sum_{j=1}^N \bar{w}_j$.

D. Numerical Results

We consider the example described in Table I to compute end-to-end delay bounds for traffic flows generated by node 1 when implementing FIFO, FP or WRR service policies, and having an associated slot $s^1 = 11$ during a cycle $c = 30$. For WRR, a heuristic weight allocation is considered based on traffic rates. The obtained upper bounds on end-to-end delays under FIFO, FP and WRR scheduling policies with the classic and extended models are described in Table II, with reference to worst-case simulation results.

c (ms)	s^1 (ms)	B (Mb/s)	$f_i(n_i, P_i, D_i, l_i, e_i)$
30	11	1	$f_1(3, 140ms, 140ms, 4K\ bits, 4ms)$ $f_2(6, 500ms, 500ms, 3K\ bits, 3ms)$

Table I. EXAMPLE PARAMETERS

	FIFO	FP		WRR	
		f_1	f_2	f_1	f_2
Classic model (ms)	87	53	87	57	117
Extended model(ms)	145	60	180	90	180
Simulation (ms)	115	59	114	61	92
Deadline (ms)	140	140	500	140	500

Table II. DELAYS WITH CLASSIC AND EXTENDED MODELS

Upper bounds on end-to end delays obtained with classic models under the different service policies infer that the system schedulability is verified. However, with reference to worst-case simulation results, we notice that classic models lead to optimistic upper bounds on end-to end delays under the different service policies. Consequently, ignoring the non-preemptive message transmission in the model may lead to wrong decision concerning system schedulability.

However, we notice that upper bounds on end-to-end delays obtained with extended models are always greater than worst-case simulation results, and infer that the system schedulability is not proved under FIFO policy, where the delay is greater than the deadline.

V. REFINED SERVICE CURVES WITH INTEGER LINEAR PROGRAMMING

To compute tighter upper bounds on delays, we refine in this section the service curves under the different policies based on ILP. First, analytical formulations of the optimization problems corresponding to the service policies are detailed. Then, the obtained parameters are integrated in the refined service curves.

A. Problem Formulation for FIFO and FP policies

Before going through the analytical formulation of the optimization problem, consider again the example in Table I under FIFO multiplexing to show the pessimism of the considered *offered TDMA time slot* allocated to node 1 in the extended model. According to Eq.7, the *lower bound of offered TDMA time slot* is $\bar{s} = \max\{s - e_1, e_2\} = 7ms$.

However, there are actually three possible scenarios to transmit messages of traffic flows f_1 and f_2 during the TDMA time slot of node 1: (i) 2 messages of flow f_1 and 1 message of flow f_2 , (ii) 2 messages of flow f_1 , (iii) 1 message of flow

f_1 and 2 messages of flow f_2 , (iv) 3 messages of flow f_2 . As we can see, the minimum *offered TDMA time slot* is actually equal to 8 ms which corresponds to scenario (ii). This value is consequently greater than the computed one with Eq.7, 7ms. Hence, we can clearly see that this computed value may lead to pessimistic delay bounds, i.e., more than the worst-case delays that may actually occur.

To obtain the minimum duration of the offered TDMA time slot, we first formulate the problem under FIFO multiplexing. Then, we extend this formulation under FP multiplexing. Our objective is to find the minimum of *offered TDMA time slot* \bar{s} .

For each end-system generating N traffic flows, we consider x_i the number of messages of traffic flow f_i that can be transmitted within a time slot s . The respective ILP problem is as follows:

minimize

$$\bar{s} = \sum_{i=1}^N x_i * e_i \quad (14)$$

subject to:

$$\sum_{i=1}^N x_i * e_i \leq s \quad (14a)$$

$$s - (\sum_{i=1}^N x_i * e_i) < e_{max} \quad (14b)$$

$$x_i \in \mathbb{N}, 1 \leq i \leq N \quad (14c)$$

where,

- the objective is to minimize the offered TDMA time slot when transmitting a variable number of messages of the different flows ($1 \leq i \leq N$) in order to maximize the remaining time and cover the worst-case scenario;
- constraint (14a) guarantees that the offered TDMA time slot \bar{s} is smaller than the allocated TDMA time slot s ;
- constraint (14b) guarantees that the remaining time with the minimum offered TDMA time slot is smaller than the maximum message delivery time e_{max} ;
- constraint (14c) guarantees that the number of transmitted messages x_i of each traffic flow f_i is non-negative integer.

The minimum offered TDMA time slot that results from this ILP problem \bar{s} is integrated in the extended service curve model described in Theorem 2 to obtain the refined service curve model detailed in the following corollary.

Corollary 1. *Let consider an end-system having the minimum offered TDMA time slot \bar{s} , a refined strict service curve guaranteed on TDMA-based network under FIFO multiplexing is*

$$\bar{\beta}(t) = \beta_{c, \bar{s}}(t - WT + (c - \bar{s}))$$

This ILP formulation can be easily extended under FP policy. To find the minimum *offered TDMA time slot* $\bar{s}_{\leq i}$ of the aggregate flow $f_{\leq i}$, we need to consider only the subset of traffic flows $\{f_1, f_2, \dots, f_i\}$ instead of all traffic flows N in the ILP problem formulation (14). The obtained minimum offered slot $\bar{s}_{\leq i}$ is then integrated in the extended service curve

described in Theorem 3 to obtain the refined service curve detailed in the following corollary.

Corollary 2. *Let consider aggregate flow $f_{\leq i}$ having the minimum offered TDMA time slot $\overline{s_{\leq i}}$, a refined strict service curve guaranteed on TDMA-based network under FP multiplexing is*

$$\overline{\beta_{\leq i}}(t) = \beta_{c, \overline{s_{\leq i}}}(t - WT_{\leq i} + (c - \overline{s_{\leq i}}))$$

Using Theorem 1 and Corollary 2, a refined service curve for flow f_i is given by:

$$\overline{\beta}_i(t) = (\overline{\beta_{\leq i}}(t) - \sum_{j=1}^{i-1} \alpha_j(t))_{\uparrow} \quad (15)$$

The optimization problem can be seen as a bin-packing problem which is known to be NP-hard. However, from a practical point of view, if the number of traffic flows is not too large (less than 100), we can solve this optimization problem efficiently in a short time.

B. Problem Formulation for WRR policy

First, we show through the example described in Table I that we can obtain tighter delay bounds if we accurately select the respective transmission time \overline{w} for each flow. Then, we try to find the best slot-distribution between the different flows to respect their stability condition and to reduce the delay bounds pessimism.

Revisiting the example, we simply transformed the respective access times $w_1 = 7.7, w_2 = 3.9$ of flows f_1 and f_2 to the refined weights $\overline{w}_1 = 4$ (1 message m_1), $\overline{w}_2 = 3$ (1 message m_2) for non-preemptive transmission. This transformation leads to the unfair treatment for flow f_1 . However, if we consider instead $\overline{w}_1 = 8$ (2 message m_1) for flow f_1 , then the delay bound becomes $\overline{D}_1 = 64$ which is lower than the one obtained with the extended model. This change shows the impact of time-access distribution between the flows on the delay bounds tightness. Hence, we will try to find the most accurate respective access time for each flow using ILP.

The ILP formulation for WRR is as follows:

$$\begin{aligned} & \text{minimize} \\ & \sum_{i=1}^N |w_i - x_i * e_i| \end{aligned} \quad (16)$$

subject to:

$$e_{max} + c - s + \sum_{i=1}^N x_i * e_i = \overline{c} \quad (16a)$$

$$\sum_{i=1}^N x_i * e_i \leq s \quad (16b)$$

$$\frac{x_i * e_i * B}{\overline{c}} \geq r_i, \forall 1 \leq i \leq N \quad (16c)$$

$$x_i \in \mathbb{N}^+, \forall 1 \leq i \leq N \quad (16d)$$

where,

- the objective function is to minimize the difference between the offered access times in preemptive and non-preemptive cases for all flows;
- constraint (16a) guarantees the respect of the maximum round robin cycle duration \overline{c} ;

- constraint (16b) guarantees that the sum of all respective access time is smaller than the initial slot duration;
- constraint (16c) is needed to guarantee the stability condition of any flow i where the offered service rate has to be greater than the traffic rate r_i .
- constraint (16d) guarantees that the number of transmitted messages x_i of each traffic flow f_i is a positive integer.

Let $(x_1^*, x_2^*, \dots, x_N^*)$ be the optimal solution of the optimization problem above. Flow f_i is allowed to transmit up to x_i^* messages during each WRR round. Hence, the respective access time is $\overline{w}_i = x_i^* * e_i$ to obtain the refined service curve defined in the following corollary.

Corollary 3. *The strict service curve guaranteed on TDMA-based network under non-preemptive WRR multiplexing is*

$$\overline{\beta}_i(t) = \beta_{c, \overline{w}_i}(t) \quad (17)$$

where $\overline{w}_i = x_i^* * e_i$ and $\overline{c} = e_{max} + c - s + \sum_{j=1}^N \overline{w}_j$

C. Numerical Results

We consider again the example in Table I and recompute the obtained delay bounds for traffic flows f_1 and f_2 under FIFO, FP and WRR policies using the refined service curve models. The obtained delay bounds using refined service curve models are described in Table III.

	FIFO	FP		WRR	
		f_1	f_2	f_1	f_2
Classic model (ms)	87	53	87	57	117
Extended model (ms)	145	60	180	90	180
Refined model (ms)	119	60	119	64	204
Simulation (ms)	115	59	114	61	92
Deadline (ms)	140	140	500	140	500

Table III. DELAY BOUNDS WITH DIFFERENT MODELS

As we can see, the delay bounds with the refined model are tighter than the extended model under FIFO, FP and WRR multiplexing and greater than worst-case simulation results. Further, the new upper bounds with refined models infer system schedulability under FIFO, which was not proved with the extended model. This fact highlights the importance of the used model and its impact on the system schedulability guarantees.

VI. VALIDATION

In this section, we evaluate the efficiency of our proposed analytical framework for timing analysis of safety-critical systems based on TDMA arbitration through a realistic avionic case study. The impact of the analytical model on the system schedulability is discussed herein with reference to simulation results.

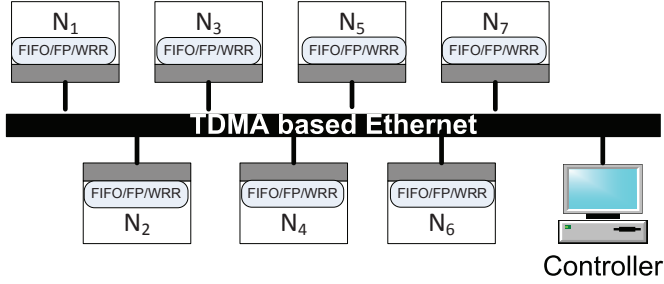
A. Case Study

Our case study is a representative avionic sensors/actuators network based on TDMA-based Ethernet at 100Mbps inter-connecting seven I/O modules and an avionic controller, as illustrated in Figure 7. We consider three scenarios where all

the nodes use either FIFO, FP or WRR. We have an equal time slots allocation among the different I/O modules where TDMA time slot duration $s = 256 \mu s$ and TDMA cycle $c = 1792 \mu s$.

There are three types of aggregate traffic flows generated by the different I/O modules and described in Table IV. Then, the numbers of messages in each aggregate traffic flow for each node are described in Table V.

Figure 7. Network Architecture



Traffic flow	Period (μs)	Deadline (μs)	delivery time (μs)
TC_1	8000	8000	60
TC_2	16000	16000	49
TC_3	32000	32000	41

Table IV. TRAFFIC FLOWS PARAMETERS

Node Ids	Traffic flow types	Number of messages
N_1, N_2	TC_1	6
	TC_2	9
N_3, N_4	TC_1	12
	TC_3	12
N_5	TC_2	17
	TC_3	47
N_6	TC_1	4
	TC_2	15
	TC_3	23
N_7	TC_1	17

Table V. GENERATED MESSAGES NUMBERS

B. Worst-case Simulation process

The worst-case simulation is performed as following. All messages of each traffic flow are the generated at the same instant, while the arrival instants of different traffic flows are randomly generated in a very small interval to capture the worst-case scenario. Under FIFO multiplexing, all messages are buffered in the same queue according to their arrival instants (First In First Out); whereas, under FP and WRR multiplexing, the messages of different traffic flows are buffered in different queues according to their traffic classes. To transmit a message, we have to check different conditions: (i) no other message is under transmission to respect the non-preemption feature; (ii) the remaining time of the TDMA slot is enough to transmit the considered message; (iii) under FP multiplexing, no higher priority message is in the queue; (iv) under WRR multiplexing, the token is assigned to its queue. The simulation results are used herein to have an idea on the tightness of the analytical delay bounds.

C. Performance Evaluation

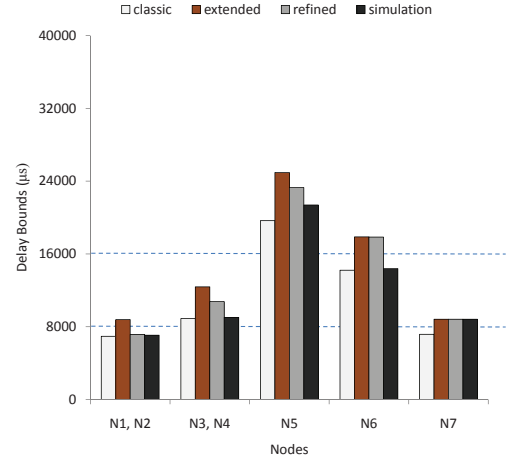


Figure 8. Delay bounds with FIFO policy

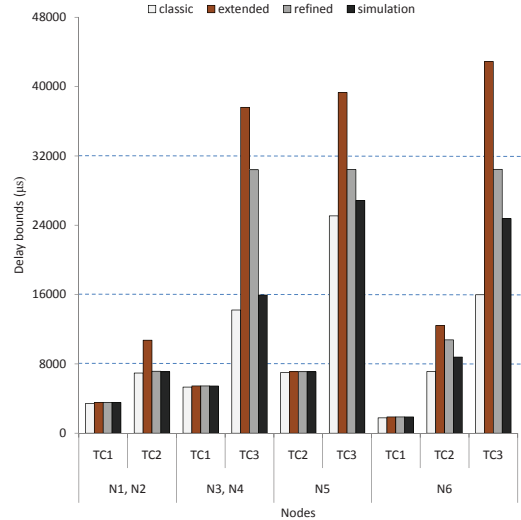


Figure 9. Delay bounds with FP policy

Comparative analysis of different models

To validate the efficiency of our proposed timing analysis, we conduct a comparative analysis of obtained delay bounds for our case study with the different service curve models detailed in the paper: classic, extended and refined models. Then, we refer to simulation results to have an idea on the tightness of obtained delay bounds.

The obtained delay bounds under FIFO, FP and WRR policies are illustrated in Figures 8, 9 and 10, respectively. These results confirm the first conclusions obtained for the example described in Table I: (i) classic models under FIFO, FP and WRR policies lead sometimes to optimistic delay bounds compared to simulation results; (ii) unlike the classic models, the extended models cover the worst-case scenario and

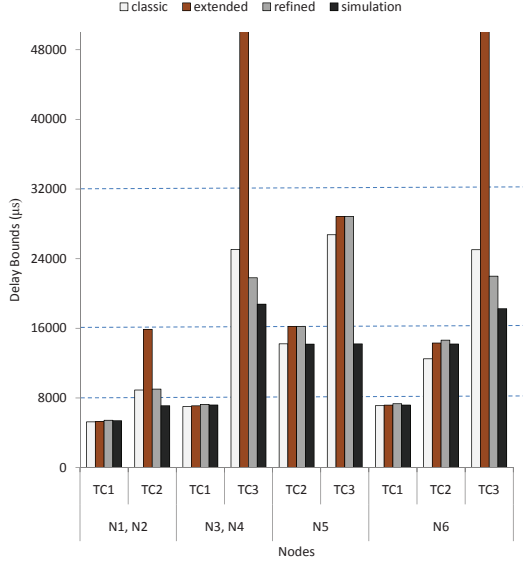


Figure 10. Delay bounds with WRR policy

provide stronger guarantees on system schedulability; however, they can sometimes lead to over-pessimistic delay bounds; (iii) refined models reduce the pessimism of delay bounds compared to extended models and consequently provide more reliable guarantees on system schedulability.

It is worth noting that the obtained delay bounds with extended and refined models are the same when we have homogeneous traffic, i.e., all the generated messages of traffic flows have the same delivery time, such as for I/O module N_7 . Furthermore, these obtained delays are the same for highest priority traffic under FP and WRR policies for the different I/O modules.

Impact of used models on System Schedulability

There are two main interesting conclusions from these obtained results concerning system schedulability when using different models.

- The first conclusion concerns the optimistic results of classic models that can lead to a wrong decision concerning the system schedulability. In fact, for the I/O module N_7 under FIFO policy, we obtain with classic model a delay bound less than the deadline which infers that the system is schedulable. However, as we can see, the obtained delay with simulation > 8 ms which means that the system is non-schedulable. Hence, the optimistic delay bounds obtained with classic models based on Network Calculus provide guarantees for messages that will actually miss their deadlines in the worst-case. This situation can result in catastrophic consequences for safety-critical systems. It is worth noting that the delay bounds obtained with the extended model for the same node and the same traffic flow is > 8 ms which infers that the system schedulability is not proved.
- The second conclusion concerns the over-pessimism of delay bounds obtained with extended models that

can also lead to the non verification of the sufficient schedulability condition, and consequently do not prove the system schedulability. In fact, for I/O modules N_1 and N_2 under FIFO policy, we obtain with extended model a delay bound greater than the deadline (8 ms) which means that the system schedulability is not proved. However, when considering the obtained delay bounds with the refined model, the sufficient schedulability condition becomes verified and the system schedulability is guaranteed. The same remark is true under FP and WRR policy for I/O modules N_3 N_4 N_6 when considering TC_3 . The delay bound of extended model is greater than the deadline 32ms and smaller with refined model. Hence, the use of refined models based on Network Calculus avoids the over-pessimism of delay bounds, and can provide stronger guarantees on the system schedulability compared to extended models.

VII. CONCLUSION

In this paper, we introduced an efficient timing analysis using Network Calculus and ILP of safety-critical systems implementing TDMA arbitration and considering the non-preemption feature impact on the obtained delay bounds under various service policies.

We first extended classic models based on a fluid flow model to a packet flow model under FIFO, FP and WRR policies. Afterwards, we refined these models based on ILP to avoid the over-pessimism of delay bounds obtained with extended models. The results for a representative avionic case study show the efficiency of our proposed models to provide stronger guarantees on system schedulability, compared to classic models.

VIII. APPENDICES

A. Appendix A

The proof of Theorem 2 is as follows.

Proof: Let $R(t), R^*(t)$ be the input and output cumulative function of the total flow, respectively. In order to prove that the obtained curve is a strict service curve, we have to prove as specified in Definition 3 that for all backlogged period $]\tau, t]$:

$$R^*(t) - R^*(\tau) \geq \bar{\beta}(t) = \beta_{c,\bar{s}}(t - \tau - WT + (c - \bar{s}))$$

We will verify Definition 3 for all the possible values of t and τ giving WT and \bar{s} , as follows:

- if $t - \tau \leq WT$, $R^*(t) - R^*(\tau) \geq 0 = \bar{\beta}(t - \tau)$
- If $WT \leq t - \tau \leq WT + \bar{s}$, let $u = \inf_w \{w \in]\tau, t] | R^*(w) < R^*(w + \varepsilon), \forall \varepsilon > 0\}$, where u is the starting time to serve flow f and $R^*(u) = R^*(\tau)$. There are two cases:

Case 1: $u > \tau$. We have $u \leq \tau + WT$, then

$$\begin{aligned} R^*(t) - R^*(\tau) &= R^*(t) - R^*(u) \\ &\geq R^*(t - (\tau + WT - u)) - R^*(u) \\ &= B(t - \tau - WT) = \bar{\beta}(t - \tau) \end{aligned}$$

Case 2: $u = \tau$. Let $v = \inf\{w \in]\tau, t[\mid R^*(w) = R^*(w + \varepsilon), \forall \varepsilon > 0\}$, where v is the ending time of a offered slot.

▷ If $t \leq v$, $R^*(t) - R^*(\tau) = B(t - \tau) > \bar{\beta}(t - \tau)$

▷ If $v < t \leq v + c - \bar{s}$,

$$\begin{aligned} R^*(t) - R^*(\tau) &\geq R^*(v) - R^*(\tau) = B(v - \tau) \\ &= \bar{\beta}(v + c + e_{max} - s - \tau) \geq \bar{\beta}(v + c - \bar{s} - \tau) \geq \bar{\beta}(t - \tau) \end{aligned}$$

▷ If $v + c - \bar{s} < t \leq \tau + WT + \bar{s}$,

let $u' = \inf\{w \in]v, t[\mid R^*(w) < R^*(w + \varepsilon), \forall \varepsilon > 0\}$, where u' is the beginning time of a slot. We have $u' \leq v + c - \bar{s}$. If $t \leq u' + \bar{s}$,

$$\begin{aligned} R^*(t) - R^*(\tau) &= R^*(t) - R^*(u') + R^*(v) - R^*(\tau) \\ &= B(t - u') + B(v - \tau) \\ &\geq B(t - (v + c - \bar{s}) + v - \tau) \\ &= B(t - \tau - (c - \bar{s})) \geq \bar{\beta}(t - \tau) \end{aligned}$$

Otherwise, $R^*(t) - R^*(\tau) \geq R^*(u' + \bar{s}) - R^*(u') = B\bar{s} \geq \bar{\beta}(t - \tau)$

- If $WT + \bar{s} < t - \tau$, there must exist t' such that $t = t' + mc$, ($m \in \mathbb{N}^+$) and $WT - (c - \bar{s}) = e_{max} - (s - \bar{s}) \leq t' - \tau < WT + \bar{s}$. Consider an arbitrary backlogged period $]t' + ic, t' + (i + 1)c[$ where $0 \leq i \leq m$ and $i \in \mathbb{N}$. The node can serve the aggregate flow at least during $c - \bar{s}$, then

$$R^*(t' + (i + 1)c) - R^*(t' + ic) \geq B\bar{s}$$

$$\begin{aligned} \text{Hence, } R^*(t) - R^*(\tau) &= R^*(t) - R^*(t') + R^*(t') - R^*(\tau) \\ &\geq Bm\bar{s} + \bar{\beta}(t' - \tau) = \bar{\beta}(t - \tau) \end{aligned}$$

Definition 3 is then verified for the different values of t and τ giving WT and \bar{s} which finishes the proof. ■

The proof of Theorem 3 is as follows.

Proof: The proof is similar to the proof of Th. 2. We have only to replace the parameters WT and \bar{s} in the first proof by the parameter $WT_{\leq i}$ and $\bar{s}_{\leq i}$, respectively. ■

B. Appendix B

The proof of Theorem 5 is as follows.

Proof: Let $R(t), R^*(t)$ be the input and output cumulative function of flow f_i , respectively. Consider an arbitrary backlogged period $] \tau, t[$, we prove that $R^*(t) - R^*(\tau) \geq \bar{\beta}_i(t - \tau)$.

- If $0 \leq t - \tau \leq \bar{c} - \bar{w}_i$, $R^*(t) - R^*(\tau) \geq 0 = \bar{\beta}_i(t - \tau)$
- If $\bar{c} - \bar{w}_i < t - \tau \leq \bar{c}$, there are 3 cases: 1) There is a complete slot s inside $] \tau, t[$. Within this slot f_i can be transmitted $\lfloor \frac{w_i}{e_i} \rfloor$ messages because the other flows can only be served in maximum amount of time $\sum_{j=1, j \neq i}^N \bar{w}_j$.

$$R^*(t) - R^*(s) \geq B * \lfloor \frac{w_i}{e_i} \rfloor * e_i = B * \bar{w}_i \geq \bar{\beta}_i(t - \tau)$$

2) There is a partial slot s inside $] \tau, t[$. The maximum time where f_i is blocked is bounded by

$$c - s + \sum_{j=1, j \neq i}^N \bar{w}_j, \text{ then}$$

$$\begin{aligned} R^*(t) - R^*(\tau) &\geq B(t - \tau - (c - s + \sum_{j=1, j \neq i}^N \bar{w}_j)) \\ &> B(t - \tau - (\bar{c} - \bar{w}_i)) = \bar{\beta}_i(t - \tau) \end{aligned}$$

3) There are 2 partial consecutive slots within $] \tau, t[$. The maximum time where f_i is blocked is bounded by $c - s + e_{max} + \sum_{j=1, j \neq i}^N \bar{w}_j$. The difference with case 2 is the impact of e_{max} which represents the worst case of idle remaining time between two consecutive slots. Hence,

$$\begin{aligned} R^*(t) - R^*(\tau) &\geq B(t - \tau - (c - s + e_{max} + \sum_{j=1, j \neq i}^N \bar{w}_j)) \\ &= B(t - \tau - (\bar{c} - \bar{w}_i)) = \bar{\beta}_i(t - \tau) \end{aligned}$$

- If $\bar{c} < t - \tau$, there must exist t' such that $t = t' + mc$, ($m \in \mathbb{N}^+$) and $0 \leq t' - \tau < c$. Hence,

$$\begin{aligned} R^*(t) - R^*(\tau) &= (R^*(t) - R^*(t - \bar{c})) \\ &\quad + (R^*(t - \bar{c}) - R^*(t - 2\bar{c})) + \dots \\ &\quad + (R^*(t') - R^*(\tau)) \\ &\geq m * B * \bar{w}_i + \bar{\beta}_i(t' - \tau) = \bar{\beta}_i(t - \tau) \end{aligned}$$

REFERENCES

- [1] A. Bouillard, L. Jouhet, E. Thierry, et al. Service curves in network calculus: dos and don'ts. 2009.
- [2] A. Bouillard and G. Stea. Exact worst-case delay for FIFO-multiplexing tandems. ValueTools, 2012.
- [3] F. Consortium et al. Flexray communications system. *Protocol Specification Version, 2*, 2005.
- [4] T. Ferrandiz, F. Frances, and C. Fraboul. A Network Calculus Model for SpaceWire Networks. RTCSA, 2011.
- [5] N. Gollan and J. Schmitt. Energy-efficient tdma design under real-time constraints in wireless sensor networks. MASCOTS, 2007.
- [6] K. Hermann, A. A., G. P., and K. Steinhammer. The time-triggered Ethernet (TTE) design. Eighth IEEE International Symposium on Object-Oriented Real-Time Distributed Computing, 2005.
- [7] A. Koubãa, M. Alves, E. Tovar, and A. Cunha. An implicit gts allocation mechanism in ieee 802.15.4 for time-sensitive wireless sensor networks: theory and practice. *Real-Time Systems*, 39(1-3).
- [8] J. Le Boudec and P. Thiran. *Network calculus: a theory of deterministic queuing systems for the internet*. springer-Verlag, 2001.
- [9] A. Mifdaoui, F. Frances, and C. Fraboul. Performance analysis of a Master/Slave switched Ethernet for military embedded applications. *IEEE Transactions on Industrial Informatics*, 6, November 2010.
- [10] E. Wandeler and L. Thiele. Optimal tdma time slot and cycle length allocation for hard real-time systems. Proceedings of the 2006 Asia and South Pacific Design Automation Conference, 2006.