IPP Hurray!

www.hurray.isep.ipp.pt

Technical Report

Energy/Delay Trade-off of the GTS Allocation Mechanism in IEEE 802.15.4 for Wireless Sensor Networks

Anis Koubaa

Mário Alves

Eduardo Tovar

Energy/Delay Trade-off of the GTS Allocation Mechanism in IEEE 802.15.4 for Wireless Sensor Networks

Anis KOUBAA, Mário ALVES, Eduardo TOVAR

IPP-HURRAY!

Polytechnic Institute of Porto (ISEP-IPP)

Rua Dr. António Bernardino de Almeida, 431

4200-072 Porto

Portugal

Tel.: +351.22.8340509, Fax: +351.22.8340509

E-mail: {akoubaa@dei, mjf, emt@dei}.isep.ipp.pt

http://www.hurray.isep.ipp.pt

Abstract

The IEEE 802.15.4 protocol proposes a flexible communication solution for Low-Rate Wireless Personal Area Networks (LR-WPAN) including wireless sensor networks (WSNs). It presents the advantage to fit different requirements of potential applications by adequately setting its parameters. When in beaconenabled mode, the protocol can provide timeliness guarantees by using its Guaranteed Time Slot (GTS) mechanism. However, power-efficiency and timeliness guarantees are often two antagonistic requirements in wireless sensor networks. The purpose of this paper is to analyze and propose a methodology for setting the relevant parameters of IEEE 802.15.4-compliant WSNs that takes into account a proper trade-off between power-efficiency and delay bound guarantees. First, we propose two accurate models of service curves for a GTS allocation as a function of the IEEE 802.15.4 parameters, using Network Calculus formalism. We then evaluate the delay bound guaranteed by a GTS allocation and express it as a function of the duty cycle. Based on the relation between the delay requirement and the duty cycle, we propose a power-efficient superframe selection method that simultaneously reduces power consumption and enables meeting the delay requirements of real-time flows allocating GTSs. The results of this work may pave the way for a powerefficient management of the GTS mechanism in an IEEE 802.15.4 cluster.

Energy/Delay Trade-off of the GTS Allocation Mechanism in IEEE 802.15.4 for Wireless Sensor Networks

Anis KOUBAA, Mário ALVES, Eduardo TOVAR

IPP-HURRAY! Research Group, Polytechnic Institute of Porto

Rua Dr. Antonio Bernardino de Almeida, 431, 4200-072 Porto, PORTUGAL

{akoubaa, emt}@dei.isep.ipp.pt , mjf@isep.ipp.pt

Abstract

The IEEE 802.15.4 protocol proposes a flexible communication solution for Low-Rate Wireless Personal Area Networks (LR-WPAN) including wireless sensor networks (WSNs). It presents the advantage to fit different requirements of potential applications by adequately setting its parameters. When in beacon-enabled mode, the protocol can provide timeliness guarantees by using its Guaranteed Time Slot (GTS) mechanism. However, power-efficiency and timeliness guarantees are often two antagonistic requirements in wireless sensor networks. The purpose of this paper is to analyze and propose a methodology for setting the relevant parameters of IEEE 802.15.4-compliant WSNs that takes into account a proper trade-off between power-efficiency and delay bound guarantees. First, we propose two accurate models of service curves for a GTS allocation as a function of the IEEE 802.15.4 parameters, using Network Calculus formalism. We then evaluate the delay bound guaranteed by a GTS allocation and express it as a function of the duty cycle. Based on the relation between the delay requirement and the duty cycle, we propose a power-efficient superframe selection method that simultaneously reduces power consumption and enables meeting the delay requirements of real-time flows allocating GTSs. The results of this work may pave the way for a power-efficient management of the GTS mechanism in an IEEE 802.15.4 cluster.

1. Introduction

With the emergence of new Wireless Sensor Network (WSN) applications under reliability and timing constraints, the provision of real-time guarantees may be more crucial than saving energy during critical situations. In such situations, it is mandatory to achieve the best trade-off between two antagonistic requirements: power-efficiency and timeliness guarantees. The IEEE 802.15.4 protocol [1] is one potential candidate to achieve predictable real-time performance for Low-Rate Wireless Personal Area Networks (LR-WPANs). We propose to analyze its performance in terms of power-efficiency and delay guarantees.

The IEEE 802.15.4 Task Group (TG4) [1], together with the ZigBee Alliance [2], has developed an entire communication protocol stack for LR-WPANs. Although the IEEE 802.15.4 protocol was not exclusively designed for wireless sensor networks, it provides suitable mechanisms for WSN applications [3]. The physical layer of the IEEE 802.15.4 protocol seems particularly suitable for WSN applications, namely in terms of data-rate, energy-efficiency and robustness. More importantly, the Medium Access Control (MAC) sub-layer, with the provision of a superframe structure bounded by two signaling beacon frames (when in beacon-enabled mode), makes the IEEE 802.15.4 protocol sufficiently flexible and able to fulfill the needs of a large variety of applications. In fact, when an IEEE 802.15.4-

compliant WPAN disables the generation of periodic beacon frames (non-beacon enabled mode), all nodes in the network compete for medium access using non-slotted CSMA/CA. The advantage of the non beacon-enabled mode, with regards to WSN application requirements, is that it easily allows scalability and self-organization. However, the non beacon-enabled mode does not provide any guarantee to deliver data frames within a certain deadline. For time-critical applications, real-time guarantees may be achieved with the beacon-enabled mode. This mode offers the possibility of allocating/deallocating time slots in a superframe, called Guaranteed Time Slots (GTSs), and providing predictable minimum service guarantees. Having a minimum service guarantee, it is possible to predict the worst-case timing performance of the network.

This paper provides a methodology, based on the Network Calculus formalism [4], for evaluating the trade-off between delay guarantees and power-efficiency for real-time applications using the GTS mechanism in one IEEE 802.15.4 cluster. The basic idea is the following. **Given a set of data flows within an IEEE 802.15.4 cluster, where each data flow has a delay requirement D, what is the most efficient network setting that satisfies the delay requirement of each data flow when it allocates one time slot GTS and minimizes the energy consumption?**

For that purpose, we propose two alternative models for the service curve provided by a GTS allocation, and derive the corresponding delay bounds. An expression of the duty cycle as a function of the delay bound is also derived. Based on this relation, we provide a methodology to meet the delay requirement of each data flow with the minimum energy consumption (lowest duty cycle) by adequately setting the IEEE 802.15.4 parameters.

To our best knowledge, this is the first work evaluating the energy/delay trade-off of the GTS allocation mechanism in IEEE 802.15.4. We hope that this work provides a first step towards the power-efficient use of the IEEE 802.15.4 GTS mechanism.

2. Related Work

Power-efficiency issues of the IEEE 802.15.4 were addressed in [13-14]. In [13], the authors analyzed the energy-efficiency of the IEEE 802.15.4 radio in a dense wireless sensor network. They have proposed an energy aware radio activation policy and analyzed the corresponding average power consumption and transmission reliability as a function of the network parameters. In [14], the authors analyzed the energy consumption of the slotted CSMA/CA MAC protocol and derived the consumed energy for successful and collided transmissions. In this paper, we focus on the analysis of the power-efficiency of the GTS allocation mechanism in IEEE 802.15.4.

From an allocation point of view, the concept of a GTS allocation is similar to a Time Division Multiple Access (TDMA) time slot allocation. A reserved amount of bandwidth is periodically granted for a given data flow. The amount of bandwidth is determined by the duration of the time slot and its periodicity. However, the IEEE 802.15.4 GTS mechanism is more flexible than a classic TDMA since the GTS duration may be dynamically adjusted by adequately setting the parameters of the IEEE 802.15.4 superframe (see Section 3), whereas a TDMA time slot duration is, in general, fixed for a given network configuration.

Moreover, TDMA and the IEEE 802.15.4 MAC approaches differ in several aspects. Actually, IEEE 802.15.4 presents several advantages as compared to TDMA for deployment in WSNs. Scalability is an important limitation of TDMA-based approaches since the number of nodes within one TDMA cluster must be kept as low as possible, which hinders its usage in WSNs. On the other hand, the IEEE 802.15.4 is expected to manage up to 254 nodes into one cluster. Another drawback of TDMA is the lack of support for dynamic topology changes (e.g. a node failure, a new node entering the network, node mobility), since the TDMA schedule has to be re-adapted to the new configuration of the network, every time the topology changes. Moreover, communications in TDMA-based networks are quite

dependent from the cluster manager. If it fails, the underlying sensor nodes will be disconnected from the entire network. This is not the case for the IEEE 802.15.4, since the protocol is able to operate either with or without a central manager and it is designed to be easily adapted to different network topologies [1].

The performance of the IEEE 802.15.4 protocol has been subject of a few research studies, however focusing more on the performance of its CSMA/CA protocol [5-7, 14] or its general characteristics using simulations [8]. In [9], the authors proposed the analysis of the GTS allocation mechanism for providing timeliness guarantees in IEEE 802.15.4 networks. While the work in [9] evaluated the impact of the IEEE 802.15.4 parameters on the guaranteed bandwidth, maximum throughput and delay bound, the power-efficiency of a GTS allocation was not addressed. This paper extends the work in [9] to capture the trade-off between energy-efficiency and delay guarantees of a GTS allocation.

In this paper, we consider the Network Calculus approach [4] since it is independent from the traffic representation and is more adapted to the computation of network delays [11]. Hence, to analyze the trade-off between the energy-efficiency and the delay bound, we derive the service curves and delay bounds guaranteed by the allocation of a GTS as functions of the protocol parameters, namely the duty cycle. The service curves and delay bounds derived in this paper are explicitly related to the parameters of IEEE 802.15.4, thus providing a suitable model for the GTS allocation scheme.

The rest of the paper is organized as follows. Section 3 gives an overview of the IEEE 802.15.4 protocol. Section 4 provides two accurate service curve models for a GTS allocation and the corresponding delay bounds. In section 5 we analyze the impact of the protocol parameters on the delay bound for a GTS allocation. Section 6 evaluates the trade-off between power-efficiency and delay guarantee as function of the IEEE 802.15.4 parameters. In Section 7 conclusions are drawn.

3. Overview of the IEEE 802.15.4 Protocol

The IEEE 802.15.4 MAC protocol supports two operational modes that may be selected by a central node called *PAN* (*Personal Area Network*) *coordinator*:

- the *non beacon-enabled mode*, in which the MAC is ruled by non-slotted CSMA/CA;
- the *beacon-enabled mode*, in which beacons are periodically sent by the PAN coordinator to identify its PAN and synchronize nodes that are associated with it.

In this paper, we only consider the beacon-enabled mode, since it has the ability of providing timeliness guarantees to the network. In beacon-enabled mode, the *Beacon Interval* (*BI*) defines the time between two consecutive beacons, and includes an active period and, optionally, an inactive period. The active period, called *superframe*, is divided into 16 equally-sized time slots, during which frame transmissions are allowed. During the inactive period (if it exists), all nodes may enter in a sleep mode, thus saving energy. Fig. 1 illustrates an example of the beacon interval and the superframe structures.
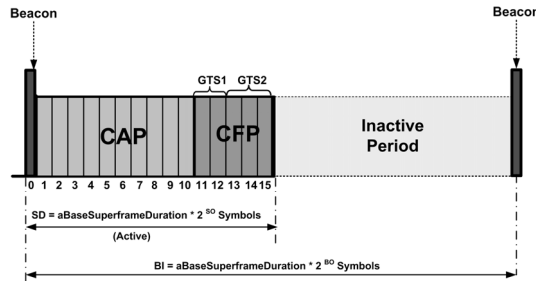

Fig. 1. Beacon Interval and Superframe Duration concepts

The Beacon Interval (*BI*) and the *Superframe Duration* (*SD*) are determined by two parameters, the *Beacon Order* (*BO*) and the *Superframe Order* (*SO*), respectively. The Beacon Interval is defined as follows:

$$BI = aBaseSuperframeDuration \cdot 2^{BO} ,$$
$$for \ \ 0 \leq BO \leq 14$$

(1)

The Superframe Duration, which determines the duration of the active period, is defined as follows:

$$SD = aBaseSuperframeDuration \cdot 2^{SO} ,$$
$$for \ \ 0 \leq SO \leq BO \leq 14$$

(2)

In Eqs. (1) and (2), *aBaseSuperframeDuration* denotes the minimum duration of the superframe, corresponding to $SO = 0$. The IEEE 802.15.4 standard fixes this duration to 960 symbols (a symbol corresponds to 4 bits). This value corresponds to 15.36 ms, assuming a 250 kbps in the 2.4 GHz frequency band. In this paper, we will only consider the features of the 2.4 GHz frequency range, which is supported by the MICAz motes[1] from Crossbow Tech. [10], for example.

By default, the nodes compete for medium access using slotted CSMA/CA within the *Contention Access Period* (CAP) during *SD*. In case of a busy channel, a node computes its backoff period based on a random number of time slots. The IEEE 802.15.4 protocol also offers the possibility of having a *Contention-Free Period* (CFP) within the superframe (Fig. 1). The CFP, being optional, is activated upon request from a node to the PAN coordinator for allocating time slots depending on the node's requirements. Upon receiving this request, the PAN coordinator checks whether there are sufficient resources and, if possible, allocates the requested time slots. These time slots are called *Guaranteed Time Slots* (GTSs) and constitute the CFP. If the available resources are not sufficient, the GTS allocation request fails. The corresponding node then must send its data frames during the CAP. A detailed description of GTS management and of the slotted CSMA/CA mechanism is presented in [3].

SD and *BI* also determine the duration of the inactive period. Saving energy may be achieved by imposing long inactivity periods leading, however, to increased communication latencies. Hence, *BO* and *SO* are key parameters for balancing the timing performance and energy savings in IEEE 802.15.4-based networks. In this paper, we show how to find the optimal durations of *SD* and *BI* that best fit the timeliness requirements (delay) of the applications, while minimizing the energy consumption.

4. Delay Bound Analysis of a GTS Allocation in an IEEE 802.15.4 Cluster

4.1 Problem Formulation

Let us consider an IEEE 802.15.4 cluster with a unique PAN coordinator, and a set of nodes within its radio coverage. The network operates in beacon-enabled mode, thus the PAN coordinator periodically sends beacon frames. The Beacon Interval and the Superframe Duration are defined by Eq. (1) and Eq. (2), respectively. Let *C* be the total data rate of the output link. In our case, the data rate is fixed to *C* = 250 kbps, as previously mentioned in Section 3.

---

[1] MICAz motes provided by Crossbow do not implement the IEEE 802.15.4 MAC protocol. They only provide the PHY layer.

We assume that each sensor node in the range of the PAN coordinator runs an application that generates a data flow. We consider that each data flow has a cumulative arrival function $R(t)$ upper bounded by the linear arrival curve $\alpha(t) = b + r \cdot t$ with b denoting the maximum burst size, and r being the average arrival rate (Fig. 2).
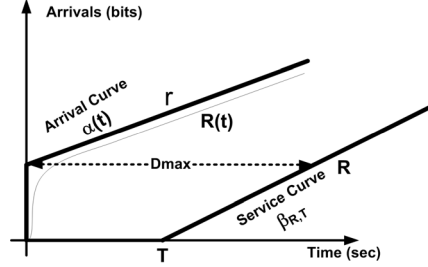

Fig. 2. Arrival curve, service curve and delay bound

This model is called a (b, r) model. In addition to its simplicity, the (b, r) model has the advantage to represent a higher bound for any kind of traffic, rather than dealing with unrealistic assumptions on the traffic type (e.g., assuming a Poisson arrival pattern). For instance, it has been shown in [11] that a periodic traffic with or without jitter can be represented with a (b,r) curve. In general, it is possible to translate any traffic type to the (b, r) model, provided that the cumulative arrival function is upper bounded. The variance between the (b, r) curve and the realistic model may be large for highly variable data rate traffic. For a periodic traffic with or without jitter, however, the variance is not too pessimistic. Since WSN applications typically generate periodic traffic, the (b, r) model is considered to be acceptable in that context.

The main purpose of our analysis is to derive an expression for the delay bound as a function of the duty cycle to evaluate the trade-off between energy consumption and delay guarantee. For that reason, we first derive the delay of a data flow $R(t)$ bounded by a (b, r) curve and that has allocated one GTS in the superframe, as a function of the IEEE 802.15.4 parameters (Section 4.3). We then generalize this result for an allocation of a GTS with n time slots (Section 4.4). Section 4.2 gives a brief introduction to the Network Calculus methodology relevant to our analysis.

4.2 Delay Bound Analysis using Network Calculus

In Network Calculus theory [4], the delay bound analysis for a given data flow with a cumulative arrival function $R(t)$ assumes the following.

1. It exists an arrival curve $\alpha(t)$ that upper bounds $R(t)$ such that $\forall s,\ 0 \leq s \leq t,\ R(t) - R(s) \leq \alpha(t - s)$. This inequality means that the amount of traffic that arrives to receive service in any interval $[s,t]$ never exceeds $\alpha(t - s)$.

2. It exists a minimum service curve $\beta(t)$ guaranteed to $R(t)$.

Then, the delay bound, D_{max}, for a data flow with an arrival curve $\alpha(t)$ that receives the service $\beta(t)$ is the maximum horizontal distance between $\alpha(t)$ and $\beta(t)$:

$$D_{\max} = h(\alpha, \beta) = \sup_{s \geq 0} \left\{ \inf \left\{ \tau \geq 0 \mid \alpha(s) \leq \beta(s + \tau) \right\} \right\} \tag{3}$$

Fig. 2 illustrates an example of the delay bound for a linear arrival curve $\alpha(t) = b + r \cdot t$ that receives a *rate-latency* service curve $\beta_{R,T}(t) = R \cdot (t - T)^+$, where $R \geq r$ is the guaranteed bandwidth, T is the maximum latency of the service and $(x)^+ = \max(0, x)$. This service curve is typically used for servers that provide a bandwidth guarantee with a certain latency. The latency T refers to the deviation of the service (e.g. blocking factor of non-preemptive transmissions).

The delay bound D_{max} (presented in Fig. 2) guaranteed for the data flow with the arrival curve $\alpha(t) = b + r \cdot t$ by the service curve $\beta_{R,T}(t) = R \cdot (t - T)^+$ is computed as [4]:

$$D_{\max} = \frac{b}{R} + T \qquad (4)$$

4.3 Delay Bound Analysis for One Time Slot GTS

In this section, we derive the delay bound guaranteed for a data flow upper bounded by an arrival curve $\alpha(t) = b + r \cdot t$ and granted one GTS slot allocation in the superframe, determined by the parameters *BO* and *SO* as defined in Eq. (1) and Eq. (2).

Based on the results of Network Calculus, our problem is reduced to finding a service curve $\beta(t)$ that corresponds to a GTS allocation. Then, it is possible to derive the expression of the delay bound using Eq. (3) and Eq. (4).

4.3.1 The GTS Service Curve

Let us consider a one time slot GTS allocated to a data flow bounded by a (b, r) curve (Fig. 3).
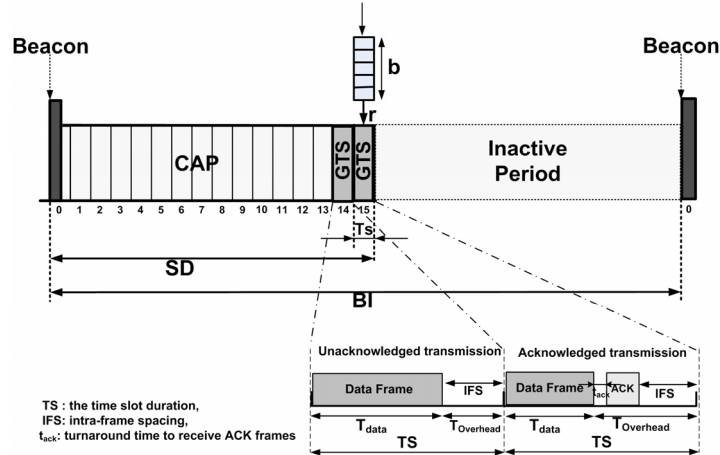

Fig. 3. The GTS service time and transmission modes

According to the IEEE 802.15.4 standard [1], a node that has allocated a GTS can transmit a message **if and only if** the whole *transaction*, including data transmission, the Intra-Frame Spacing (IFS) and the acknowledgement (if requested), can be completed before the end of the GTS. Otherwise, it must wait until the next GTS in the next superframe. Fig. 3 also presents the cases of acknowledged and unacknowledged transactions. Hence, a given GTS allocation may impose a restriction on the frame length. Moreover, only a part of the GTS can be used for data

transmission. The rest will be idle or used by a potential acknowledgement frame. The impact of a given GTS allocation on the guaranteed bandwidth will be discussed later on, in Section 5.

For the sake of simplicity and without loss of generality, we assume one data frame transmission in a GTS. However, note that these results are still valid from multiple data frame transmissions inside a GTS.

Now, let Ts denote the Time Slot duration in the superframe. Then,

$$Ts = \frac{SD}{16} = abaseSuperframeDuration * 2^{(SO-4)} \tag{5}$$

We define T_{data} and T_{idle} (Fig. 3) such that:

$$Ts = T_{data} + T_{idle} \tag{6}$$

T_{data} defines the **maximum duration** used for data frame transmission inside a GTS. T_{idle} is the sum of *idle times* spent inside a GTS due to protocol overheads (IFS and/or Ack frames). The minimum value of T_{idle} comprises the time required for an IFS, T_{IFS}, and a potential acknowledgement in the case of an acknowledged transmission, T_{ack}. Hence, the following analysis holds for both acknowledged and unacknowledged transactions. The only difference is the setting of the value of T_{idle}. We can write:

$$T_{idle} = T_{Overhead} + T_{wasted}$$
$$where \ T_{Overhead} = T_{IFS} + T_{Ack} \cdot \amalg_{Ack} \tag{7}$$

where $\amalg_{Ack} = 1$ for an acknowledged transaction and $\amalg_{Ack} = 0$ for an unacknowledged transaction. The value of *wasted time, T_{wasted}* is greater than zero if the length of a GTS is longer than the transaction time (transmission + IFS + Ack).

A frame of the data flow that has allocated the GTS uses the entire capacity of the output link for a time period of T_{data}, every Beacon Interval *BI*.

a. Real service curve of a GTS

First, we need to define the maximum latency T that a burst may wait for a service. This latency occurs for a burst that arrives just after the end of the GTS. It results that the maximum latency T (Fig. 4) can be computed as:
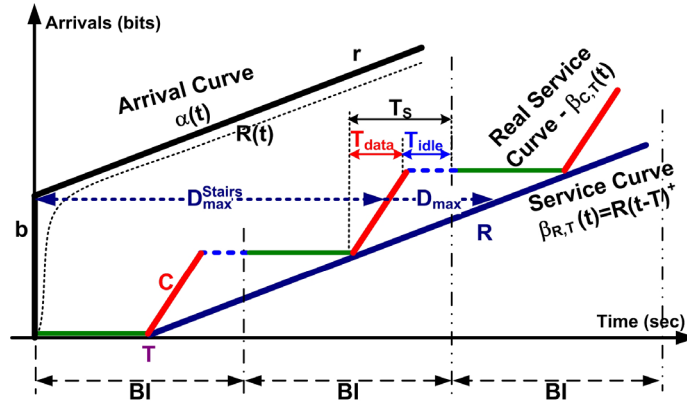
$$T = BI - Ts \tag{8}$$

Fig. 4. The GTS service curves

In our analysis, we define the kth *superframe* as the superframe contained in the time interval $\left[(k-1)\cdot BI, \ k\cdot BI\right]$.

In the first superframe, the service curve received by the data flow; that is, the minimum number of bits that must be transmitted during the GTS, is:

$$\beta_{C,T}^1(t) = \begin{cases} C\cdot\left(t-\left(BI-Ts\right)\right)^+ & \forall t, \ 0 \leq t \leq BI -T_{idle} \\ 0 & \text{Otherwise} \end{cases}$$

where $(x)^+ = \max(0,x)$. In the second superframe, the service curve received by the data flow is:

$$\beta_{C,T}^2(t) = \begin{cases} C\cdot T_{data} + C\cdot\left(t-\left(2\cdot BI-Ts\right)\right)^+ \\ \qquad\qquad \forall t, \ BI \leq t \leq 2\cdot BI -T_{idle} \\ 0 \qquad\qquad \text{Otherwise} \end{cases}$$

By recurrence, the general expression that gives the service curve received in the kth superframe is:

$$\beta_{C,T}^k(t) = \begin{cases} (k-1)\cdot C\cdot T_{data} + C\cdot\left(t-\left(k\cdot BI-Ts\right)\right)^+ \\ \qquad \forall t, \ (k-1)\cdot BI \leq t \leq k\cdot BI -T_{idle} \\ 0 \qquad\qquad \text{Otherwise} \end{cases} \tag{9}$$

As a result, the overall service curve $\beta_{C,T}^{stair}(t)$ is defined as:

$$\beta_{C,T}^{stair}(t) = \sum_k \beta_{C,T}^k(t) \qquad \forall t \tag{10}$$

$\beta_{C,T}^{stair}(t)$ is a stair function, which is the sum of all rate-latency service curves of each superframe, and represents an overall service curve for the data flow (Fig. 4).

b. Approximation to a rate-latency service curve

Hereafter, we show that the stair service curve $\beta_{C,T}^{stair}(t)$ can be simplified to a basic rate-latency service curve $\beta_{R,T}(t)$, where R is the guaranteed bandwidth of one time slot GTS.

In fact, R is the amount of bits sent in each beacon interval divided by BI, and is expressed as follows:

$$R = \frac{T_{data}\cdot C}{BI} = \left(\frac{Ts-T_{idle}}{BI}\right)\cdot C \tag{11}$$

We denote DC as the duty cycle, which is defined as:

$$DC = \left(\frac{SD}{BI}\right) = 2^{IO} \leq 1 \quad \text{where} \quad IO = SO - BO \tag{12}$$

IO represents the *Inactivity Order*. Then, applying Eqs. (5) (11) and (12), the guaranteed bandwidth can be rewritten as follows:

$$R = \frac{2^{IO}C}{16} - \frac{T_{idle}}{BI}C = \lambda\cdot DC\cdot C - w_{idle} \tag{13}$$

$$\text{where } w_{idle} = w_{overhead} + w_{wasted}$$

with $\lambda = 1/16$, and w_{idle} representing the amount of *unused bandwidth* into a GTS due to protocol overheads (IFS and/or Ack frame) and wasted time inside the GTS.

Eq. (13) represents the guaranteed bandwidth as a function of the duty cycle. As a consequence, since R is the guaranteed bandwidth of one time slot GTS and $T = BI - Ts$ is the maximum latency, then $\beta_{R,T}(t) = R \cdot (t - T)^+$ is a rate-latency service curve that corresponds to one time slot allocation. In Fig. 4, observe that $\beta_{R,T}(t)$ is a good approximation of $\beta_{C,T}^{stair}(t)$.

In summary: A GTS allocated in a superframe defined by a beacon order *BO* and a superframe order *SO* provides a minimum service curve $\beta_{C,T}^{stair}(t)$ expressed in Eq. (10), which may be approximated by a rate-latency service curve $\beta_{R,T}(t) = R \cdot (t - T)^+$ with an guaranteed bandwidth R expressed in Eq.(13) and a latency T expressed in Eq. (8).

4.3.2   Delay Bound

a.   Case of a rate-latency service curve

Based on the results of Eq. (4), and if we consider the approximation of the service curve $\beta_{R,T}(t)$, the delay bound experienced by a data flow with an arrival curve $\alpha(t) = b + r \cdot t$, which has allocated one time slot GTS, can be computed as follows:

$$D_{max} = \frac{b}{\lambda \cdot DC \cdot C - w_{idle}} + (BI - Ts) \tag{14}$$

b.   Case of a stair service curve

A more accurate delay bound is obtained by considering the stair service curve $\beta_{C,T}^{stair}(t)$. The delay bound is represented by the maximum horizontal deviation $h\left(\alpha, \beta_{C,T}^{stair}\right)$ between the arrival curve $\alpha(t) = b + r \cdot t$ and $\beta_{C,T}^{stair}(t)$. The maximum horizontal deviation is reached at the angular point defined by the intersection between the slope of $\alpha(t)$ and the y-axis, so at the burst size b (Fig. 4).

Let us consider a burst size $b \le C \cdot T_{data}$. Then, the delay bound is the maximum horizontal deviation between the angular point and the first stair of $\beta_{C,T}^{stair}(t)$; that is, $\beta_{C,T}^1(t)$. In this case, the delay bound D_{max}^{stair} is:

$$D_{max}^{stair} = \frac{b}{C} + (BI - Ts) \qquad \text{if } b \le C \cdot T_{data}$$

Now, if $C \cdot T_{data} < b \le 2 \cdot C \cdot T_{data}$ then, the delay bound is the maximum horizontal deviation between the angular point and the second stair function of $\beta_{C,T}^{stair}(t)$; that is, $\beta_{C,T}^2(t)$. This example is shown in Fig. 4. In this case, the delay bound D_{max}^{stair} is:

$$D_{max}^{stair} = \frac{b - C \cdot T_{data}}{C} + 2 \cdot BI - Ts$$
$$= \frac{b}{C} + 2 \cdot BI - Ts - T_{data}$$
$$\text{if } C \cdot T_{data} < b \le 2 \cdot C \cdot T_{data}$$

In general, when $k \cdot C \cdot T_{data} < b \le (k+1) \cdot C \cdot T_{data}$, we show that the delay bound of a data flow with an arrival curve $\alpha(t) = b + r \cdot t$, which has allocated one time slot GTS is:

$$D_{\max}^{stair} = \frac{b}{C} + (k+1) \cdot BI - Ts - k \cdot T_{data}$$
$$\text{if } k \cdot C \cdot T_{data} < b \le (k+1) \cdot C \cdot T_{data} \tag{15}$$

4.4 General Case of n Time Slot GTS

The reasoning for a GTS with n allocated time slots is similar to 1-allocated time slot except that the service will not be continuous due to the mandatory idle period inside each slot for IFS and acknowledgement processing. This causes a slight change in the construction of the stair service curve for an allocation of n time slots when $2 \le n \le 7$ (no more than seven GTSs can be allocated in each superframe). Fig. 5 shows an example for allocations of 1, 2 and 3 time slots inside a superframe.

Fig. 5. The GTS service curve for n time slot allocation

As a consequence, the guaranteed bandwidth of a GTS with n time slots is:

$$R_n = n \cdot (\lambda \cdot DC \cdot C - w_{idle}) \tag{16}$$

and the latency T is then:

$$T_n = BI - n \cdot TS \tag{17}$$

So, the delay bound of a data flow with (b,r) curve by the service curve $\beta_{R,T}(t)$ becomes:

$$D_{n,\max} = \frac{b}{n \cdot (\lambda \cdot DC \cdot C - w_{idle})} + (BI - n \cdot Ts) \tag{18}$$

A more precise bound is obtained with the $\beta_{C,T}^{stair}(t)$ service curve as follows:

$$D_{n,\max}^{stair} = \left(\frac{b}{C} + (k+1) \cdot BI - n \cdot (Ts + k \cdot T_{data}) + m \cdot T_{idle} \right)$$
$$\text{where } m = \left\lfloor \frac{b - k \cdot (n \cdot T_{data}) \cdot C}{T_{data} \cdot C} \right\rfloor \tag{19}$$
$$\text{if } k \cdot C \cdot (n \cdot T_{data}) < b \le (k+1) \cdot C \cdot (n \cdot T_{data})$$

Note that this relationship between the delay and the guaranteed bandwidth is different from that of other guaranteed rate servers such as Weighted Fair Queuing (WFQ) [12]. This is because in WFQ and similar servers the guaranteed bandwidth R and the latency T are completely independent. In such systems, the latency is fixed. However, in the IEEE 802.15.4 GTS mechanism, the bandwidth and the latency are correlated, since they both depend on the superframe parameters BO and SO. Hence, increasing the number of slots allocated to a GTS simultaneously increases the guaranteed bandwidth and reduces the maximum latency, leading to a lower delay bound.

5. Impact of the Beacon and Superframe Orders on the Delay Bound

In this section, we investigate the impact of the beacon and superframe orders on the delay bound. The problem is to determine the best configuration of the superframe structure (SO) that reduces the delay bound for a given duty cycle.

Referring to Eq. (14), the delay bound typically depends on the guaranteed bandwidth R and on the beacon interval (service latency). Observe that in Eq. (11) the guaranteed rate depends on the portion of time slot T_{data} used for effective data transmission (without taking into account the overheads (IFS and acknowledgements)), which also depends on the frame size (and its corresponding IFS) and on the superframe order. In [9], the authors have evaluated the maximum bandwidth guaranteed for one time slot GTS as a function of frame size and their corresponding IFSs. Fig. 6 illustrates the guaranteed bandwidth per one GTS allocation for different superframe orders.
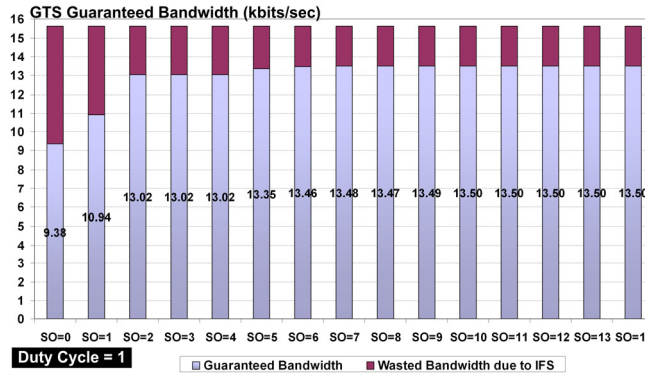

Fig. 6. Guaranteed bandwidth (kbps) per one GTS allocation

We observe that the guaranteed bandwidth is slightly above 13 kbps, except for low superframe orders. The guaranteed bandwidth for $SO = 0$ and $SO = 1$ is relatively low compared to the others. This is due to the impact of IFS, since the time slot durations are too small to send high amounts of data.

Fig. 7 shows the delay bound of a GTS allocation as a function of the burst size for a 100% duty cycle and an average arrival rate of 5 kbps.

Observe that the superframe order that provides the lowest delay bound depends from the burst size. For low burst size values (0.5 kbits, 1 kbit), the delay bound increases as a function of the superframe order. In this case, the impact of the latency (BI-Ts) on the delay bound is more important than the impact of the guaranteed bandwidth, which turns the superframe order $SO = 0$ the most suitable for providing the lowest delay bound.

However, for higher burst size values (2, 4, 7, and 10 kbits), the delay bound as a function of the superframe order is no longer monotonic. Actually, lower delays are achieved for superframe orders higher than 0. For example, for a burst size of 10 kbits, the lowest delay bound is guaranteed with a superframe order $SO = 2$. Moreover, note that in this case the delay bounds for superframe orders $SO = 3$ and $SO = 4$ are lower than those guaranteed with $SO = 0$. This

behavior is explained by the impact of the guaranteed bandwidth on the delay bound. In fact, when the burst size is relatively high, the impact of the guaranteed bandwidth R is more important than the latency ($BI - Ts$) especially for the 0-2 superframe orders. The low guaranteed bandwidth for $SO = 0$ explains well a higher delay bound than in $SO \in \{1, 2, 3, 4\}$, for a burst size equal to 10 kbits.
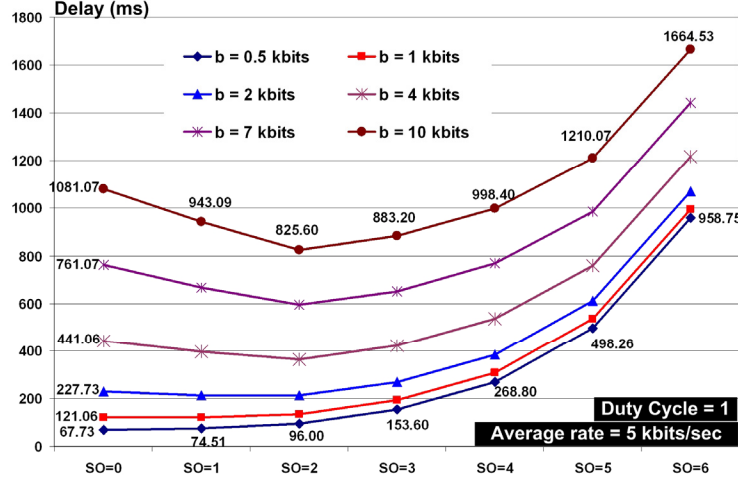

Fig. 7. Delay bound (ms)

In summary, for common WSN applications with low burst sizes and low data rates, it is more likely that superframe order $SO = 0$ is the most suitable for providing real-time guarantees thanks to its reduced latency. This is at the cost of lower throughput and guaranteed bandwidth, but this is not important for low arrival curves. However, if the burst size is relatively high (more than 1 kbit), the superframe order $SO = 2$ is better for providing timeliness guarantees thanks to its higher guaranteed bandwidth.

6. Power-Efficiency of a GTS Allocation

6.1 Problem Statement

At start-up, the PAN coordinator of a given WPAN cluster must choose a superframe structure. The choice of the superframe structure affects the timing performance as well as the energy consumption in the cluster. Saving energy requires superframe structures with low duty cycles, whereas improving the timing performance requires higher duty cycles. Therefore, a trade-off must be achieved by choosing the lowest duty cycle that still satisfies the timing constraints of the data flow. In this study, we consider data flows with a common arrival curve and the same delay requirement. This assumption can hold for many wireless sensor network applications, since the data sent by sensor nodes are usually related to common observed phenomena, and thus data flows are closely similar. In case the arrival curves are different from one node to another, we consider the data flow with the highest burst size and the lowest delay requirement for the worst-case dimensioning of the system.

In this section, we investigate the following question.

Given a set of data flows within a PAN cluster, where each data flow has a maximum burst size b and a per-hop delay requirement D, what is the most efficient superframe structure, i.e. the combination of SO and BO, that satisfies the delay requirement of each data flow when it allocates one time slot GTS and minimizes the energy consumption?

6.2 Duty Cycle Evaluation as a Function of the Delay Bound

The purpose of this section is to derive the expression of the duty cycle as a function of the delay bound to evaluate the trade-off between energy consumption and delay guarantee.

Note that each data flow allocating a GTS with the above requirements must adapt its arrival rate to the amount of bandwidth guaranteed by its GTS, corresponding to the combination of *SO* and *BO*, based on Eq. (14).

Stating that the combination of *SO* and *BO* simultaneously satisfies the delay requirement *D* and reduces the energy consumption is equivalent to state that the corresponding duty cycle $DC=2^{SO-BO}$ is the lowest one that satisfies the delay requirement *D*. In what follows, we envisage expressing the duty cycle as a function of the delay.

Based on Eqs. (14) and (11), we have:

$$D = \frac{b}{\frac{T_{data}^{full}}{BI} \cdot C} + (BI - Ts)$$

Taking into account Eqs. (2) and (12), then the duty cycle can be expressed as a function of the delay as follows:

$$DC = \frac{SD}{D + \lambda \cdot SD} \cdot \left(\frac{b}{T_{data} \cdot C} + 1 \right) \tag{20}$$

where *SD* is the Superframe Duration (Eq. (2)).

According to the standard, the minimum valid value of the duty cycle is then:

$$DC = 2^{IO} \quad \text{where } IO = \left\lfloor \log_2 \left(\frac{SD}{D + \lambda \cdot SD} \cdot \left(\frac{b}{T_{data} \cdot C} + 1 \right) \right) + 1 \right\rfloor \tag{21}$$

Based on Eq. (21), the duty cycle depends on three parameters: the burst size, the delay bound and the superframe order. Note that T_{data} also depends on the superframe order (refer to Section 4.3.1).

6.3 Duty Cycle versus Superframe Order

A first problem is to determine the adequate superframe order reducing the duty cycle and still meeting a delay bound *D* for a given burst *b*. Fig. 8 shows the variation of the duty cycle as a function of the superframe order for different burst size values. The delay bound requirement is assumed to be 3 seconds.

Observe that for relatively low burst sizes (0.1 kbits, 1 kbits) the minimum duty cycle required to satisfy the delay bound increases with the superframe order.

For a burst size equal to 10 kbits, there is no advantage of using low superframe orders for $SO \in \{0, 1, 2, 3, 4\}$. The duty cycle remains the same, since lower superframe orders have lower service curve latencies whereas higher superframe orders provide higher guaranteed bandwidths.
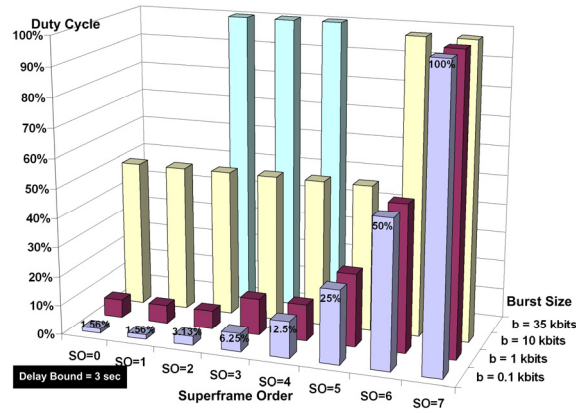
Fig. 8. Duty cycle versus Superframe Order

However, for a burst size $b = 35$ kbits, only superframe orders $SO \in \{2, 3, 4\}$ can satisfy the delay bound of 3 sec, with a full duty cycle. This is because, as stated in the previous section, the guaranteed bandwidth has the most important impact on the delay bound. Even with a full duty cycle, $SO \in \{0,1\}$ cannot satisfy the delay bound. Hence, for high burst sizes, it possible that lowest superframe orders $SO \in \{0,1\}$ cannot satisfy a delay bound with a given duty cycle while the subsequent superframe orders can.

For wireless sensor networks with low burst size, superframe order $SO = 0$ is the most suitable to reduce the duty cycle, i.e. energy consumption, for providing a given delay bound.

6.4 Duty cycle versus Delay Bound

Another problem is to determine the impact of the delay bound on the duty cycle for a given superframe order SO and a given burst b. Fig. 9 shows the variation of the duty cycle as a function of the delay bound for different superframe orders. The burst size is equal to 200 bits.
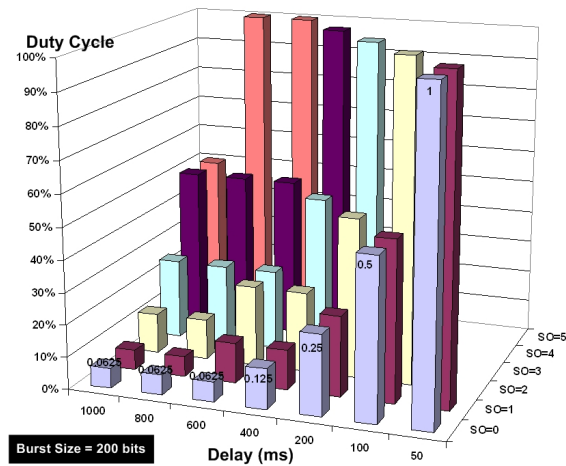

Fig. 9. Duty Cycle versus delay bound

Observe in Fig. 9 that decreasing the delay requirement does not automatically increase the duty cycle. For instance, delay values in the [600, 1000] ms range have the same 6.25% duty cycle for $SO = 0$. This fact is due to the slotted behavior of the superframe structure (defined in Eqs. (1) and (2)). Hence, in some cases, relaxing the delay requirement will not automatically lead to a lower duty cycle for some IEEE 802.15.4 superframes. It is also observed in Fig. 9 that the number of possible superframe structure configurations (alternatives for BO and SO) increases with the delay. Hence, for low delay requirements, only the lower superframe orders (*for low burst size*) can meet these delays (if possible) due to the increased latency for large SO values (that also leads to greater BO values).

7. Conclusions

In this paper, we have addressed the power-efficiency versus delay guarantee trade-off in IEEE 802.15.4-compliant wireless sensor network under timing constraints. We considered the analysis of the GTS mechanism of the IEEE 802.15.4 protocol for time-sensitive applications and proposed a methodology to select the lowest duty cycle that enables meeting the delay requirements of data flows allocating GTSs. Using Network Calculus, we have proposed two accurate models for the service curve provided by a GTS allocation and expressed the delay bound as a function of the duty cycle, which captures the relation between delay guarantee and energy consumption.

Our results show that, in case of wireless sensor networks with low burst size, lower superframe orders ($SO = 0, 1, 2$) are the most suitable for providing lower delay bound with the minimum duty cycle. However, it has been shown that the duty cycle does not necessarily decrease when increasing the delay requirement due to the slotted behavior of the protocol. In case of high burst sizes, it may happen that delay guarantee is not monotonically increasing with the superframe orders. In fact, we have shown that for a burst size of 35 kbits, only $SO = 2, 3$ and 4 are able to meet the delay requirements of the flows at a full duty cycle. This is typically due to the impact of the guaranteed bandwidth on the delay bound.

Even though the results presented in this paper are valid for star-based networks, the proposed model of the IEEE 802.15.4 GTS mechanism is general enough to be applied to multi-hop IEEE 802.15.4/Zigbee WSNs. In fact, a natural extension of this work can consider cluster-tree topologies supported by the IEEE 802.15.4/Zigbee protocols. The Zigbee standard proposes a simple beacon frame scheduling strategy for achieving multiple cluster synchronization, avoiding beacon frame collisions between multiple coordinators (which are also cluster-heads). The idea is that each coordinator sends its beacon frames and triggers its superframe duration, while the other clusters are in sleep mode. Since each coordinator can allocate GTSs to the nodes associated to it (in its cluster), our GTS model can be used to derive the service curve corresponding to each GTS in a given cluster. Based on advanced Network Calculus theorems, it is then possible to evaluate the end-to-end delay bounds and to predict the resource requirements at intermediate coordinators in multiple cluster WSNs.

8. Acknowledgement

References

[1]   IEEE 802.15.4 Standard-2003, "Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs)", IEEE-SA Standards Board, 2003.

[2] ZigBee Alliance, ZigBee Specification – June 2005, http://www.caba.org/standard/zigbee.html

[3] A. Koubâa, M. Alves, and E. Tovar, "IEEE 802.15.4 for Wireless Sensor Networks: A Technical Overview", IPP-HURRAY Technical Report (TR-050702), Jul. 2005.

[4] J-Y. Leboudec, and P. Thiran, *A Theory of Deterministic Queuing Systems for the Internet*, LNCS, Vol. 2050, P. 274 , ISBN: 3-540-42184-X, 2001.

[5] J. Mišic, and V. B. Mišic, "Duty Cycle Management in Sensor Networks Based on 802.15.4 Beacon Enabled MAC", Ad Hoc and Sensor Wireless Networks Journal, Old City Publishing, 1(3):207-233, 2005.

[6] J. Mišic and V. B. Mišic, "Access Delay and Throughput for Uplink Transmissions in IEEE 802.15.4 PAN", Elsevier Computer Communications Journal, 28(10):1152-1166, June 2005.

[7] J. Mišic, S. Shafi, and V. B. Mišic, "The Impact of MAC Parameters on the Performance of 802.15.4 PAN", Elsevier Ad hoc Networks Journal, 3(5):509–528, 2005.

[8] J. Zheng, and M. L. Lee, "A Comprehensive Performance Study of IEEE 802.15.4", IEEE Press Book, 2004.

[9] A. Koubâa, M. Alves, and E. Tovar , "GTS Allocation Analysis in IEEE 802.15.4 for Real-Time Wireless Sensor Networks", to appear in 14th International Workshop on Parallel and Distributed Real-Time Systems (WPDRTS 2006), invited paper in the special track on Wireless Sensor Networks, Apr. 2006.

[10] MICAz DataSheet , http://www.xbow.com/

[11] A. Koubâa, and Y. Q. Song, "Evaluation and Improvement of Response Time Bounds for Real-Time Applications under Non-Preemptive Fixed Priority Scheduling", International Journal of Production and Research (IJPR), (Taylor & Francis), Vol. 42, No. 14, pp. 2899-2913, July 2004.

[12] A. Demers, S. Keshavt, and S. Shenker, "Analysis and Simulation of Fair Queuing Algorithm", In Proc. of the ACM Computer Communication Review (SigComm89), pp. 3– 12, 1989.

[13] B. Bougard et al. "Energy Efficiency of the IEEE 802.15.4 Standard in Dense Wireless Microsensor Networks: Modeling and Improvement Perspectives" in Proc. Of Design, Automation, and Test in Europe Conference (DATE'05), Munich, (Germany), 7-11 March, 2005.

[14] T. Park, T. Kim, J. Choi, and S. Choi and WH Kwon, "Throughput and energy consumption. Analysis of IEEE 802.15.4 slotted CSMA/CA", IEEE Electronics Letters, vol. 41, issue 18, pp. 1017-1019, Sept. 2005.