

# Analysis of the delay of confirmed downlink frames in Class B of LoRaWAN

François Delobel, Nancy Rachkidy, Alexandre Guitton

► **To cite this version:**

François Delobel, Nancy Rachkidy, Alexandre Guitton. Analysis of the delay of confirmed downlink frames in Class B of LoRaWAN. 85th Vehicular Technology Conference (VTC), Jun 2017, Sydney, Australia. IEEE Vehicular Technology Conference (VTC). <hal-01471673>

**HAL Id: hal-01471673**

**<https://hal-clermont-univ.archives-ouvertes.fr/hal-01471673>**

Submitted on 20 Feb 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Analysis of the delay of confirmed downlink frames in Class B of LoRaWAN

François Delobel<sup>(1,2)</sup>, Nancy El Rachkidy<sup>(3,2)</sup>, Alexandre Guitton<sup>(3,2)</sup>

(1) Clermont Université, Université d’Auvergne, BP 10448, F-63000 Clermont-Ferrand, France

(2) CNRS, UMR 6158, LIMOS, F-63175 Aubière, France

(3) Clermont Université, Université Blaise Pascal, BP 10448, F-63000 Clermont-Ferrand, France

Emails: francois.delobel@udamail.fr, {nancy.el\_rachkidy,alexandre.guitton}@univ-bpclermont.fr

**Abstract**—Wireless sensor networks have long been characterized by short range communications. Lately, long range wireless sensor networks such as LoRaWAN (long range wide area network) have emerged, increasing the communication range at the cost of reduced bit rate. This paper focuses on the MAC layer of LoRaWAN and studies communications from a gateway to an end-device. We propose a Markov chain model to estimate the delivery delay of an acknowledged communication (with retransmissions) and study the impact of the number of channels, the data-rate and the number of nodes.

## I. INTRODUCTION

LPWANs (low-power wide area networks) are based on recent radio technologies for long range, low-power communications, such as LoRa [1], Sigfox [2] and Ingenu [3]. LPWAN technologies have started a new trend for Internet of Things (IoT) and monitoring applications. Indeed, unlike IEEE 802.15.4 [4], these long range technologies can efficiently cover wide areas such as whole cities [5], forests, etc.

LoRaWAN [6] specifies a MAC layer designed for the physical layer LoRa. Recently, researchers have started to study LoRaWAN, which is the only public standard for LPWANs.

General discussions on LoRaWAN can be found in [7], [8], [9]. In [7], the authors explain how LPWAN technologies can achieve ubiquitous IoT connectivity. In [8], the authors survey long range cellular networks for machine-to-machine communications, which includes LoRa. In [9], the authors discuss some security issues of LoRaWAN. A detailed discussion on LoRa is given in [10]: several of our hypotheses are based on this analysis. The only paper giving a performance study of LoRaWAN is [11]: the authors focus on the computation of the expected delay for the activation procedure.

LoRaWAN builds a star topology where low-power end-devices communicate with a network server through gateways. LoRaWAN defines three classes of communication: Class A, Class B and Class C. Class A enables low-power communications, initiated by end-devices. Class B enables periodic, low-power communications from the gateway to end-devices, called *downlink communications*. Class C enables frequent downlink communications, but requires high energy consumption for the end-devices.

In this paper, we study the delay for downlink communications. We focus on Class B of LoRaWAN because Class B is optimized for energy-efficient downlink communications.

Our contributions are three-fold: (i) we analyze in details the Class B specification, and we highlight some limitations, (ii) we propose an analytic model based on a Markov chain for downlink communications with retransmissions, and (iii) we compute the expected delay of these communications under a variety of parameters.

The remainder of this paper is as follows. Section II describes the LoRaWAN standard. Section III presents some issues of Class B and introduces our analytic model. Section IV shows our simulation settings and results. Finally, Section V concludes our work.

## II. DESCRIPTION OF LORAWAN

The three classes of LoRaWAN [6] are described in the following. Note that Class B and Class C include all the mechanisms of Class A.

In Class A, communications are always initiated by an end-device. When an end-device has data to send, it randomly selects an available sub-band and channel, and starts the transmission. Each transmission of an end-device is followed by two reception slots, which allow downlink communications from the gateway. These two reception slots are called *receive windows*, and are denoted RX1 and RX2. The channel used by the gateway for RX1 is the same as the uplink channel. The channel used by the gateway for RX2 is a predefined channel. The receive windows are the only possibility for downlink communications in Class A. After each transmission, a device cannot transmit again on the same sub-band for a long duration (typically 99 or 999 times the transmission duration, depending on the duty-cycle for this sub-band). This duty-cycle allows end-devices to sleep most of the time, and guarantees compliance with the regulations on medium usage. However, the gateway has to wait for an uplink communication in order to send downlink frames.

In Class B, the gateway broadcasts a periodic beacon every 128s. Between two beacons, each end-device opens  $N$  reception slots called *ping slots*. Downlink communications can occur during these ping slots, as well as during the reception slots that follow a transmission (according to Class A). With this mechanism, Class B enables downlink communications with bounded delay. The ping period is the delay between two ping slots, and is constant. The offset between the beacon and the first ping slot is pseudo-randomly chosen after each

beacon by each end-device. Finally, the beacon transmission is protected by a reserved period of 5.12s, which includes a guard time of 3s.

Figure 1 gives an example of Class B. The periodic beacons from the gateway are denoted  $b_1$  and  $b_2$ . According to the Class A feature, the end-device initiates the transmission of a data frame  $d_1$  at an arbitrary time. After the transmission, the end-device opens the two receive windows, and refrains from sending on this channel until the time-off duration has expired. According to the Class B feature, the gateway sends a data frame  $d_2$  at the beginning of the fourth ping period. The end-device detects the preamble during the ping slot, and remains active until it has received the whole frame.

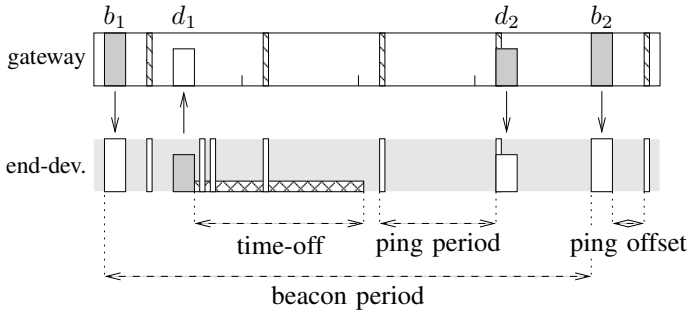


Figure 1. Class B of LoRaWAN. Dark rectangles represent transmissions, and white rectangles represent receptions.

In Class C, each end-device is continuously listening for potential frames from the gateway. Class C enables low latency downlink communications, at the cost of a very high energy consumption. In the following, we do not consider Class C.

Confirmed downlink transmissions are implemented using an acknowledgment (ACK) and retransmission mechanism: when an end-device receives a confirmed data frame from the gateway, it replies with an ACK. If no ACK is received, the gateway retransmits the data frame. Note that the standard does not specify the ACK timeout nor the number of retransmissions of the data frame by the gateway.

### III. DISCUSSION AND PROPOSITION

In this section, we first discuss some issues of Class B. Then, we present our Markov model. Finally, we describe how to compute the expected delay from the Markov model.

#### A. Discussion on limitations of Class B

We believe that there are issues in Class B in the current version of the standard (version 1.0 [6]). We discuss these issues in the following.

1) *Duty-cycle limitation for the gateway*: It is not possible for the gateway to send ACKs for a large number of confirmed uplinks. The gateway cannot acknowledge more than one frame every 66.3 seconds for each sub-band (for DR0 and with an ACK payload of 3 bytes). The gateway cannot acknowledge frames for three or more end-devices per sub-band, with saturated end-devices and a duty-cycle of 1% (for DR0 and for a payload of 30 bytes).

It is not possible for the gateway to use most ping slots of Class B. Indeed, each time the gateway sends a small data frame with 10 bytes of payload on a ping slot, it has to wait for 98.13 seconds (with DR0) before using the next ping slot (for all slots using this sub-band). Thus, the number of ping slots  $N$  impacts the delay for the first downlink communication, but has a limited impact on the downlink throughput.

In this paper, we consider that the gateway can acknowledge all data frames, and that it can use all ping slots if necessary.

2) *Conflicts between Class A and Class B*: The standard does not forbid an end-device to transmit during the beacon transmission of the gateway (as Class A end-devices are not aware of the timing of the beacons) or during the ping slots.

In this paper, we assumed that Class B end-devices refrain from sending during the beacon reserved time and ping slots.

3) *Delay before ACKs and sub-band availability*: The standard does not specify the delay between a confirmed downlink frame from the gateway and the ACK from the end-device (unlike the random backoff during the transmission of a confirmed uplink frame and the reception of the ACK from the gateway, as the ACK is supposed to be received during RX1 or RX2). Moreover, the standard does not specify the delay between the retransmissions of the same frame.

In practice, we can consider that the gateway is able to estimate the sub-band availability of each end-device, by tracking all frames sent by each end-device. When the gateway receives a data frame, it can compute the time-off duration for the end-device, based on the data-rate and on the frame length. When a data frame is not received by the gateway, the gateway assumes that the sub-band is available for the end-device: in this case, the gateway under-estimates the sub-band availability. However, it is not possible for the gateway to over-estimate the sub-band availability: it is not possible for an end-device to send a frame while the gateway assumes that the sub-band is unavailable.

In this paper, we considered that the ACK transmission from an end-device occurs as soon as possible (that is, immediately after the sub-band becomes available). After waiting for one symbol, if no transmission is detected, the gateway initiates the retransmission procedure: it retransmits the data frame during either the next ping slot or the next receive window.

Figure 2 shows an example of several transmissions. First, an end-device sends a frame  $d_1$  which is received by the gateway. Then, the gateway sends a frame  $d_2$  during a ping slot. The gateway knows that it cannot receive the ACK until the beginning of the third ping period and initiates the timeout accordingly. When the timeout expires, which is depicted by the dotted line, the gateway expects to receive the ACK. However,  $a_1$  is dropped. Then, the gateway resends the frame  $d_3$  at the beginning of the next ping slot. The end-device is unable to send another ACK, and has to wait. In the meanwhile, the gateway keeps retransmitting the frame (see frame  $d_4$ ), due to the under-estimation of the sub-band availability. Finally, the end-device sends the ACK  $a_2$  during the last ping period. It is important to notice that during the timeout (which occurred during the whole second ping period),

it is not possible for the end-device to send a data-frame, as all sub-bands were busy. Thus, the gateway either retransmits the frame at the ping slot that follows the timeout (which was the case for frame  $d_3$  on this example), or it is possible that the end-device sends a data frame between the end of the timeout and the beginning of the next ping slot.

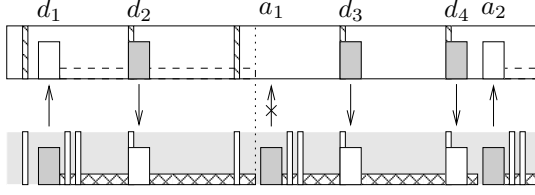


Figure 2. When the gateway estimates the timeout for an ACK from an end-device, it is not possible for the gateway to receive another transmission from this end-device during the timeout.

The standard also mentions that RX2 uses a predefined channel. If RX2 is shared by all end-devices and is used by the gateway for one end-device, it cannot be used again until the sub-band becomes available. Note that RX2 is generally used by the gateway when the transmitting end-device chose for RX1 a sub-band which was not available for the gateway.

In this paper, we assumed that RX2 can be used whenever needed. In other words, we assume that either it uses a sub-band with a large duty-cycle, or that it is on a different sub-band for each end-device.

### B. Analytic model

We make the three following hypotheses: (i) there is no queuing delay: the gateway has a single frame to send, (ii) uplink and downlink communications do not interfere with each other (due to the use of an opposite chirp [10]), and (iii) we ignore the capture effect: if two end-devices transmit simultaneously on the same channel, both frames are dropped.

Figure 3 shows our Markov chain model for confirmed downlink communications in Class B of LoRaWAN. The Ready state corresponds to the generation of a frame by the gateway. The Beacon state represents the beacon reserved period. The PSlot $_i$  states (for  $i \in [1; N]$ ) represents the  $i$ -th ping slot, while PSlot $_{N+1}$  represents a virtual ping slot where no transmission is possible (it is used only for the delay computation). The PWait $_1$  state represents the random duration between the beacon and the first ping slot. The PWait $_i$  state, for  $i \in [2; N]$ , represents the ping period between the  $(i-1)$ -th ping slot and the  $i$ -th ping slot. The PWait $_{N+1}$  state represents the (random) duration between the  $N$ -th ping slot and the beacon. The Data $_i^1$  (for  $i \in [1; N]$ ) represents the transmission of data during the  $i$ -th ping slot (that is, Class B transmission). The Data $_i^2$  (for  $i \in [1; N+1]$ ) represents the transmission of data during the  $i$ -th ping period (that is, during the receive window of a Class A transmission). The ACK state, which is final, represents the reception of an ACK from the end-device. The NoAck $_i^j$  states indicates that no ACK was received after state Data $_i^j$ , and that a retransmission will occur.

The Ready state is the initial state of our Markov chain. Then, the chain goes to the Beacon state, if the frame was generated during the beacon reserved period, or to a PWait state, if the frame was generated during a ping period. If the frame is generated during the beacon reserved period, the chain goes to the PWait $_1$  state. Let us now focus on this first ping period, represented by the dashed rectangle in the middle of the figure, as it is similar to all other periods (except the  $(N+1)$ -th period). Two cases can occur: either the gateway waits for the ping slot to occur, which corresponds to PSlot $_1$ , or a Class A transmission occurs during this ping period, which corresponds to the Data $_1^2$  state. After PSlot $_1$ , the data is sent in the Data $_1^1$  state and is either acknowledged (Ack state) or not (NoAck $_1^1$  state). After Data $_1^2$ , the data is either acknowledged (Ack state) or not. When there is no ACK, the chain goes to a PWait $_j$  state, possibly several ping periods later, depending on the retransmission delay. Let us now focus on the  $(N+1)$ -th ping period, represented by the dashed rectangle on the left. There is no ping slot, so the PSlot $_{N+1}$  state goes directly to the beacon reserved period (without Data $_{N+1}^1$  state).

Let  $B$  be the beacon period and  $N$  the number of ping slots. Ping period  $P$  is  $(B - 5.120)/N$ . Each end-device transmits with probability  $\tau \in [0; 0.01 \cdot n_{sb}]$ .  $\alpha$  is the channel quality. There are  $n_A$  other activated end-devices in the network. Let us denote by  $q_A = 1 - \tau/(n_c \cdot n_{sb})$  the probability that an end-device does not transmit on a given channel.

Probabilities  $p_{r,b}$  and  $p_{r,w_i}$  (for  $i \in [1; N+1]$ ) are the probability that the data is generated during the beacon transmission or during the  $i$ -th ping period. Those probabilities depend on the duration of each period:  $p_{r,b}$  is equal to the probability that the data is generated during the beacon reserved period (including the guard time),  $p_{r,w_i}$  is the same for all  $i \in [2; N]$ , and  $p_{r,w_1}$  and  $p_{r,w_{N+1}}$  depend on the expected value of the ping offset, which is chosen randomly within  $[0; P]$ .

$$p_{r,b} = 5.120/B, \quad (1)$$

$$p_{r,w_1} = p_{r,w_{N+1}} = P/(2B), \quad (2)$$

$$\forall i \in [2; N], p_{r,w_i} = P/B. \quad (3)$$

Probability  $p_{w_1,d_1^2}$  is the probability that the end-device sends a data frame during the ping offset period (which lasts on average  $P/2$ ), and that this frame is correctly decoded by the gateway. Probabilities  $p_{w_i,d_i^2}$ , for  $i \in [2; N]$ , and  $p_{w_{N+1},d_{N+1}^2}$  are similar, except that the corresponding ping periods last respectively  $P$  and  $P - P/2$  on average.

$$p_{w_1,d_1^2} = p_{w_{N+1},d_{N+1}^2} = \alpha \cdot \tau \cdot P/2, \quad (4)$$

$$p_{w_1,s_1} = p_{w_{N+1},s_{N+1}} = 1 - p_{w_1,d_1^2}, \quad (5)$$

$$\forall i \in [2; N], p_{w_i,d_i^2} = \alpha \cdot \tau \cdot P, \quad (6)$$

$$\forall i \in [2; N], p_{w_i,s_i} = 1 - p_{w_i,d_i^2}. \quad (7)$$

Probability  $p_{d_i^1,a}$ , for  $i \in [1; N]$ , is the probability that the end-device has correctly received the data frame from the gateway (modeled by  $\alpha$ ), that there is no interference with the

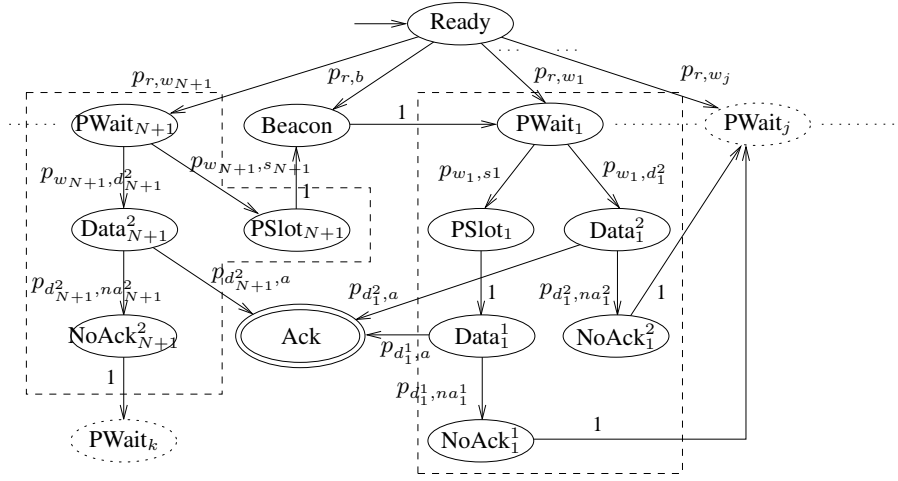


Figure 3. Markov chain model for Gateway using class B.

ACK sent by the end-device, and that the ACK is correctly decoded by the gateway.

$$\forall i \in [1; N], p_{d_i^1,a} = \alpha \cdot q_A^{n_A} \cdot \alpha, \quad (8)$$

$$\forall i \in [1; N], p_{d_i^1,na_i^1} = 1 - p_{d_i^1,a}. \quad (9)$$

Probability  $p_{na_i^1,w_j}$ , for  $i \in [1; N]$ , is equal to one for a given  $j$  due to the sub-band availability under-estimation. The value of  $j$  depends on the value of  $i$  and of the duration of the timeout. It is computed in Subsection III-C.

Probability  $p_{d_i^2,a}$ , for  $i \in [1; N+1]$ , is the probability that the end-device has correctly received the data frame (modeled by  $\alpha$ ), that there is no interference with the ACK sent, and that the ACK is correctly decoded by the gateway.

$$\forall i \in [1; N+1], p_{d_i^2,a} = \alpha \cdot q_A^{n_A} \cdot \alpha, \quad (10)$$

$$\forall i \in [1; N+1], p_{d_i^2,na_i^2} = 1 - p_{d_i^2,a}. \quad (11)$$

Figure 4 represents the three cases where the gateway sent the data frame  $d_2$  during a receive window of a frame  $d_1$  and has to retransmit the frame. Case 1 is when the ACK is not received, which yields to a sub-band under-estimation. In this case, the gateway keeps resending the data frame at each ping slot ( $d'_2$  and  $d''_2$ ), until it receives the ACK  $a'_2$ . Case 2 is when  $d_2$  is not received by the end-device, and the end-device has no other data frame to send. In this case, the gateway uses a ping slot to retransmit the data frame. Case 3 is when the  $d_2$  is not received by the end-device, and the end-device has another data frame  $d_3$  to send. In this case, the gateway retransmits the data frame during RX1 or RX2, and waits for the sub-band to become available for the end-device in order to receive  $a_2$ . All these cases correspond to different timings.

Probability  $p_{na_i^2,w_j}$ , for all  $i \in [1; N+1]$ , is equal to one: after the timeout for the ACK, the gateway is in a ping period, and the retransmission can occur either during the ping slot (see Cases 1 and 2) or a receive window (see Case 3).

### C. Computation of expected delay

Recall that the time when the ACK should be received depends on the sub-band availability. This duration depends

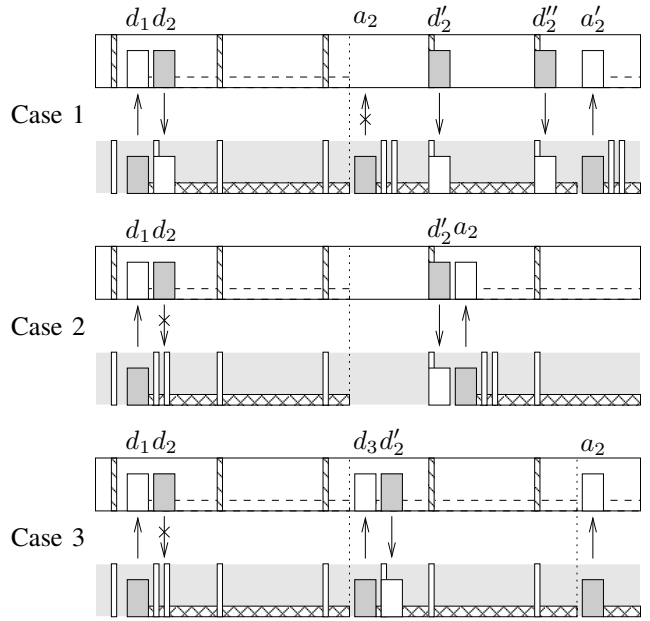


Figure 4. When the gateway has to retransmit a data frame due to a frame loss, three cases can occur, leading to different timings.

on the network saturation: if  $\tau = 0.01 \cdot n_{sb}$ , the network is saturated, and the expected delay for the sub-band availability is  $t_{off}/2$ . Otherwise, the delay is  $d_{timeout} = p_{off} \cdot t_{off}/2$ , where  $p_{off} = 1 - (0.01 - \tau/n_{sb})/0.01$  and  $t_{off} = 99 \cdot t_{frame}$ . After this duration, the Markov chain is in the  $j$ -th ping period (by neglecting the beacon reserved periods), where both values  $j = 1$  and  $j = N + 1$  are equally likely with probability 0.5 if  $(i + \lfloor d_{timeout}/P + 1/2 \rfloor) \bmod N = 0$  (that is, the first and the last ping periods), or  $j = 1 + ((i + \lfloor d_{timeout}/P + 1/2 \rfloor) \bmod N)$  with probability one otherwise.

The expected delay of the confirmed downlink can be computed from our Markov chain model with  $I \cdot \mathcal{N} \cdot D$ , where  $I$  represents the initial state of the chain,  $\mathcal{N} = (1 - Q)^{-1}$  is the

fundamental matrix of our absorbing chain,  $Q$  is our matrix without the absorbing states, and  $D$  is the column-vector with the delay of each state.

The delays of  $D$  for each state are the following.  $D_r = 0$ .  $D_b = 5.120$  as it is the duration of the beacon reserved period (including the guard time).  $D_a = d_{ack}$ , where  $d_{ack}$  is the time on air of the ACK frame.  $D_{w_i} = 0$  for all  $i \in [1; N + 1]$ , as we decided to take into account the delay either in the  $s_i$  state, or in the  $d_i^2$  state. This is due to the fact that the Markov chain can reach a  $Wait_i$  state either at the beginning of ping period  $i$  (after the beacon period for instance), or at an arbitrary time during ping period  $i$  (after a timeout).  $D_{s_1} = D_{s_{N+1}} = P/4$  as the duration of the first and the last ping periods is  $P/2$ .  $D_{s_i} = P/2$ , for  $i \in [2; N]$  as the duration of the other ping periods is  $P$ .  $D_{d_i^1} = d_{frame} + d_{subband1}$  for all  $i \in [1; N]$ : indeed, the duration corresponds to the time on air of the data frame plus the average remaining time before a sub-band becomes available for the end-device.  $d_{subband1} = d_{timeout}$ . We also have  $D_{d_1^2} = D_{d_{N+1}^2} = P/4 + d_{frame} + d_{subband2}$  and  $D_{d_i^2} = P/2 + d_{frame} + d_{subband2}$ , for  $i \in [2; N]$ , where  $d_{subband2}$  is the average remaining time before a sub-band becomes available for the end-device, given that the end-device just sent a data frame. The only difference between  $d_{subband1}$  and  $d_{subband2}$  is that we know that one of the sub-bands is busy in this case. If  $n_{sb} = 1$ ,  $d_{subband2} = t_{off} - 1 - d_{ack}$  (if we assume that RX1 is used by the gateway). Otherwise,  $d_{subband2} = d_{subband1}$ . Note that  $P/4$  or  $P/2$  is added here as we included in this state the delay of the ping period.  $D_{na_i^1} = sym$ , for  $i \in [1; N]$ , and  $D_{na_i^2} = sym$ , for  $i \in [1; N + 1]$ , where  $sym$  is the symbol duration. This delay is used by the gateway to detect if the ACK is sent or not. Also note that in our model,  $d_{subband1}$  and  $d_{subband2}$  do not depend on the number of sub-bands: this is because we consider that each sub-band is homogeneously saturated (due to the fact that  $\tau$  depends on  $n_{sb}$ ).

#### IV. PERFORMANCE EVALUATION

In this section, we present our simulation results.

We focused on the EU863-870MHz ISM band regional settings of the standard. ETSI regulations [12] for the ISM band specify the following four sub-bands (for a bandwidth of 125kHz): sub-band G with 15 channels and 1% duty-cycle, sub-band G1 with 3 channels and 1% duty-cycle, sub-band G2 with 2 channels 0.1% duty-cycle, sub-band G3 with 1 channel and 10% duty-cycle, and sub-band G4 with 1 channel and 1% duty-cycle. In the following, we assumed that all sub-bands have a duty-cycle of 1%. We varied the number of sub-bands  $n_{sb}$  from 1 to 3, and the number of channels per sub-band  $n_c$  from 1 to 3. Given the fact that the gateway has no energy limitation and that its confirmed frames are likely to be important, we assumed that there is no limitation of the number of retransmissions. The link quality is set to  $\alpha = 0.99$ .

##### A. Delay as a function of the data-rate

Figure 5 and Figure 6 show the delay of confirmed downlink frames as a function of the data-rate, for respectively  $N = 2$

and  $N = 4$  ping periods. We assumed here that there are  $n_A = 10$  competing nodes.

Figure 5 shows that the data-rate has a large impact on the delay, as expected, since increasing the data-rate reduces the time on air of frames, and thus the time-off duration. The number of sub-bands  $n_{sb}$  also has a significant impact: the larger  $n_{sb}$ , the smaller the delay, as there are more opportunities to transmit a frame. Increasing the number of channels has a limited impact on the delay, however.

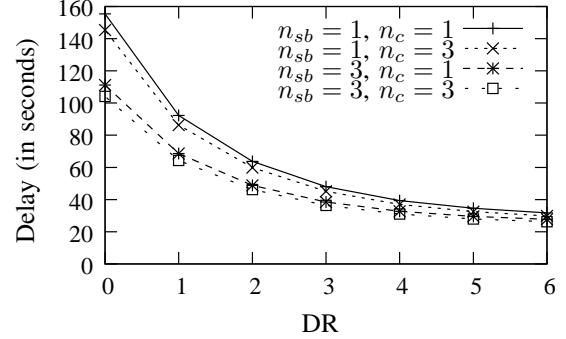


Figure 5. Delay as a function of the data-rate, for  $N = 2$  ping periods. Increasing the data-rate and  $n_{sb}$  significantly reduces the delay.

Figure 6 shows that the data-rate has a large impact on the delay. However, both  $n_{sb}$  and  $n_c$  have a limited impact. By comparison with Figure 5, it can be seen that increasing the number of ping periods from  $N = 2$  to  $N = 4$  significantly reduced the delay.

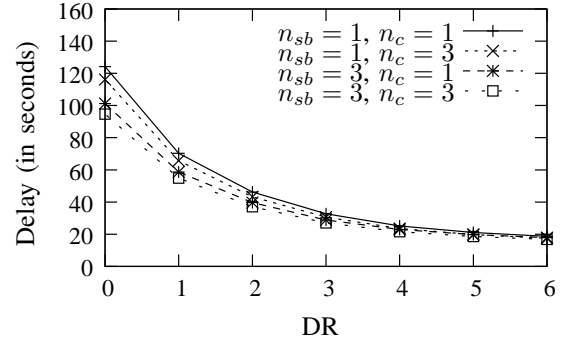


Figure 6. Delay as a function of the data-rate, for  $N = 4$  ping periods. Increasing the data-rate significantly reduces the delay.

##### B. Delay as a function of other nodes

Figure 7 and Figure 8 show the delay of confirmed downlink frames as a function of the number of competing nodes, for respectively  $N = 2$  and  $N = 4$  ping periods. We assumed here that the data-rate is DR0.

Figure 7 shows that the number of competing nodes greatly impacts the delay, especially when the number of channels is low. Indeed, all competing nodes are likely to send frames on the same channel, which results into collisions and retransmissions. Increasing the number of sub-bands from 1 to 3 results into a delay reduction which varies between 25% and 30%.

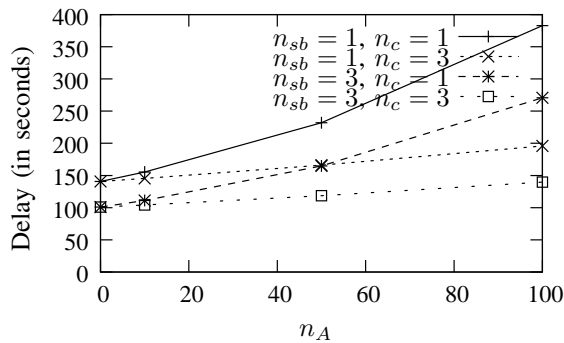


Figure 7. Delay as a function of the number of competing nodes, for  $N = 2$  ping periods. Increasing the number of competing nodes or reducing the number of sub-bands significantly increases the delay.

Figure 8 shows similar results as Figure 7. Again, increasing the number of sub-bands from 1 to 3 results into a delay reduction which varies between 18% and 22%. By comparison with Figure 7, it can be shown that increasing  $N$  significantly reduces the delay.

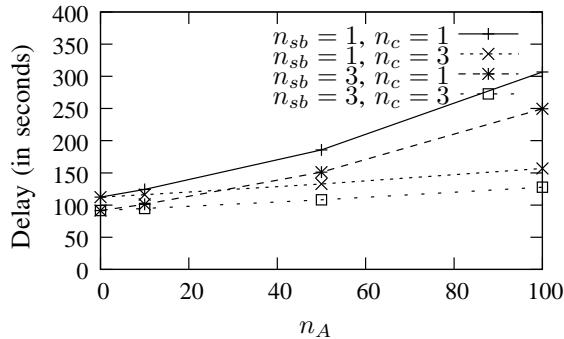


Figure 8. Delay as a function of the number of competing nodes, for  $N = 4$  ping periods. Increasing the number of competing nodes or reducing the number of sub-bands significantly increases the delay.

### C. Summary

In practice, the choice of  $N$  is a trade-off between the energy consumption of end-devices (as they have to wake up shortly at each ping period) and the delay of downlink frames. It is likely that in many Class B scenarios,  $N$  will be larger than 4, which guarantees that a downlink frame can be sent in less than 32s. Thus, according to Figure 6 and Figure 8, we expect  $n_{sb}$  and  $n_c$  to have a limited impact on the delay of Class B downlink frames, while the data-rate and  $n_A$  have a significant impact on the delay.

## V. CONCLUSION

LoRaWAN is a recent MAC standard from the industrial community for long range, low-power communications. In this paper, we focus on the delay of frames that have to be acknowledged by end-devices. Consequently, we focus on Class B of the standard, which enables downlink communications with bounded delay. We first discuss some limitations of the current version of the standard. Then, we propose a

Markov chain model for the delay. Our simulation results show that the data-rate and the number of sub-bands have a significant impact on the delay. Moreover, when the number of competing nodes is large, a large number of channels significantly decreases the delay.

**Acknowledgment:** This work has been partially funded by the Clervolc laboratory of excellence.

## REFERENCES

- [1] LoRa Alliance, <https://www.lora-alliance.org/>.
- [2] Sigfox, <http://www.sigfox.com>.
- [3] Ingenu, <http://www.ingenu.com>.
- [4] IEEE 802.15, "Part 15.4: low-rate wireless personal area networks (LR-WPANs)," IEEE, Standard for local and metropolitan area networks IEEE Std 802.15.4-2011, 2011.
- [5] M. Centenaro, L. Vangelista, A. Zanella, and M. Zorzi, "Long-range communications in unlicensed bands: the rising stars in the IoT and smart city scenarios," *IEEE Wireless Communications*, 2016.
- [6] N. Sornin, M. Luis, T. Eirich, T. Kramp, and O. Hersent, "LoRaWAN Specification," LoRa Alliance, Standard V1.0, 2015.
- [7] L. Vangelista, A. Zanella, and M. Zorzi, *Long-Range IoT Technologies: The Dawn of LoRa™*. Springer International Publishing, 2015, pp. 51–58.
- [8] Q. Song, L. Nuaymi, and X. Lagrange, "Survey of radio resource management issues and proposals for energy-efficient cellular networks that will cover billions of machines," *EURASIP Journal on Wireless Communications and Network*, pp. 1–20, 2016.
- [9] G. Margelis, R. Piechocki, D. Kaleshi, and P. Thomas, "Low throughput networks for the IoT: Lessons learned from industrial implementations," in *IEEE World Forum on Internet of Things (WF-IoT)*, 2015, pp. 181–186.
- [10] C. Goursaud and J.-M. Gorce, "Dedicated networks for IoT: PHY / MAC state of the art and challenges," *EAI endorsed Transactions on Internet of Things*, 2015.
- [11] J. Toussaint, N. El Rachkidy, and A. Guitton, "Performance analysis of the on-the-air activation in LoRaWAN," in *IEEE Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 2016.
- [12] ETSI, "Electromagnetic compatibility and radio spectrum matters (ERM); short range devices (SRD); radio equipment to be used in the 25 MHz to 1000 MHz frequency range with power levels ranging up to 500 mW; part 1: Technical characteristics and test methods," ETSI, EN 300 220-1 v2.4.1, 05 2012.