A discrete-time queueing model with periodically scheduled arrival and departure slots

Johan van Leeuwaarden[1], Dee Denteneer[2], and Jacques Resing[3]

[1] EURANDOM
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

[2] Digital Signal Processing Group
Philips Research
5656 AA Eindhoven, The Netherlands

[3] Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

Abstract

We consider a time-slotted queueing model where each time slot can either be an arrival slot, in which new packets arrive, or a departure slot, in which packets are transmitted and hence depart from the queue. The slot scheduling strategy we consider describes periodically, and for a fixed number of time slots, which slots are arrival and departure slots. We consider a static and a dynamic strategy. For both strategies, we obtain expressions for the probability generating function of the steady-state queue length and the packet delay. The model is motivated by cable-access networks, which are often regulated by a request-grant procedure in which actual data transmission is preceded by a reservation procedure. Time slots can then either be used for reservation or for data transmission.

1 Introduction

We consider a time-slotted queueing model where each time slot can either be an arrival slot, in which new packets can arrive, or a departure slot, in which packets can be transmitted and hence depart from the queue. The decision whether a time slot is an arrival or a departure slot is periodically made for a fixed number of time slots. Observe that the arrival process in this model depends on the service process, which differs from conventional assumptions on periodically structured queueing models (e.g. Van Eenige [13], Norimatsu et al. [9], Kang & Steyaert [12]).

The model is motivated by data transmission procedures in cable networks. These networks are often regulated by a request-grant mechanism in which actual data transmission is preceded by a contention-based reservation procedure to prevent collisions of data packets. Both the reservation procedure and the actual data transmission take place on the same communication channel. Hence, each time slot (corresponding to the time needed to send a 64 byte data packet on the channel) can either be used for the reservation procedure or for the data transmission. Our discrete-time queue models the size of the data queue, i.e. the number of data packets for which transmission has already been requested, but that are still waiting to be transmitted. Clearly, if a time slot is used for reservation, new packets can enter this queue (arrival slot), and if a time slot is used for data transmission, a packet can leave this queue (departure slot).

Cable networks are characterised by a substantial data transmission delay, due to which scheduling decisions must be taken in advance, so that they can be communicated to the stations. Similarly, reservations must travel from the station to the scheduler, again causing delay. Consequently, there is a time lag between a reservation and the corresponding data transmission and this time lag must consist of a number of time slots that is greater than or equal to the round-trip time in the network. Thus one is naturally led to consider periodical scheduling, for which time slots are grouped together into frames composed of both reservation and transmission slots. The nature of each slot in the frame is periodically determined and broadcast to all the stations. The timing is such that each station is aware of the layout of a frame before it actually starts.

In this paper we consider two possible scheduling strategies. The first strategy uses no information about the system's state and leads to a *fixed boundary model*. Within a frame, a fixed number of arrival slots are scheduled first, and the remaining slots are departure slots. Clearly, if there is no data to transmit during a departure slot, this capacity is lost. Therefore, the second strategy considered, is one that uses the unused departure slots as arrival slots. This flexible strategy leads to a *flexible boundary model*, which emphasizes the fact that the division of a frame into arrival and departure slots can vary from one frame to another.

Intuitively, a system implementing the flexible strategy is more efficient than a system implementing the fixed strategy. Yet, a system designer will want to have a clear quantitative understanding of the benefits of the flexible strategy, in order to offset it against the costs. It is the purpose of this paper to provide such understanding by analysing the delay in either model. For this, we will first consider the backlog: the size of the data queue at the frame boundaries. The probability generating function of the stationary distribution of the backlog in either model can be determined following a classical procedure from Bailey [2]. Due to the inclusion of the periodically scheduled reservation slots however, it is more complicated to analyse the stationary distribution of the delay in such networks. Here, we will use the techniques developed in Bruneel & Kim [4] and Kang & Steyaert [12].

In Denteneer et al. [5] bounds on the mean queue length and mean packet delay have been derived for the fixed boundary model. In Sala et al. [11] the flexible boundary model has been investigated through simulation, which shows that inducing arrival slots at the beginning of each frame might positively influence the system's performance. It is therefore that we investigate the same kind of strategy. The fixed and flexible boundary model are examples of queueing models with periodic service. In Van Eenige [13] these have been applied to traffic lights and logistic systems, and in Norimatsu et al. [9] to an IEEE 1394 serial bus, where in both publications it is assumed that the arrival process is continuous. In Kang & Steyaert [12] a queue with both periodic service and a correlated arrival process is investigated. As mentioned, in the present paper the arrival process stops during departure slots, a feature that is also captured by Resing & Örmeci [10] when considering a tandem queue with coupled processors.

The remainder of this paper is structured as follows. In Section 2, we describe and motivate the models in more detail. In Section 3, we analyse the stationary distribution of the backlog; the fixed and flexible boundary model are analysed in Section 3.1 and 3.2, respectively. A closed form expression for the probability generating function of the packet delay for both models is derived in Section 4. A numerical comparison of the two strategies is given in Section 5, followed by some suggestions for future research in Section 6.

2 Models

In this section we first introduce both the fixed and flexible boundary model. Next, we discuss the model assumptions in view of the cable network application.

Time is assumed to be slotted, with a given slot duration. In case of the fixed boundary model the schedule of each frame is fixed. That is, a frame defined as f consecutive slots consists of c arrival slots followed by $s := f - c$ departure slots. Let the random variable Y_{ti} denote the number of arriving packets during the ith arrival slot of frame t, and assume that the sequence Y_{ti}, $t = 1, 2, \ldots$ is i.i.d.. We further assume that packets that arrive during frame t cannot depart from the queue until the beginning of frame $t + 1$. This leads one to consider the following recursion

$$X_{t+1} = (X_t - s)^+ + \sum_{i=1}^{c} Y_{ti}, \tag{1}$$

where X_t denotes the backlog at the beginning of frame t and $x^+ := \max(0, x)$.

The fixed boundary model seems wasteful, in the sense that if the backlog is smaller than s, it leaves time slots unused which could alternatively be scheduled as arrival slots. This motivates the flexible boundary model in which these unused time slots are designated as arrival slots, yielding the following recursion

$$\tilde{X}_{t+1} = (\tilde{X}_t - s)^+ + \sum_{i=1}^{c+(s-\tilde{X}_t)^+} Y_{ti}, \tag{2}$$

where, for notational purposes, we add a wiggle to the random variables related to the flexible boundary model. We refer to the c arrival slots that are scheduled at the beginning of every frame as forced arrival slots.

We now comment on the model assumptions for both (1) and (2) in view of the cable network application. As mentioned, the round-trip delay inherent to such cable networks causes one to consider frame based scheduling, see e.g. Golmie et al. [6, 7]. Usually, the frame length, f, and the number of forced arrival slots, c, are chosen so that s is greater than or equal to the round-trip time, d say. Thus, one ensures that a schedule for frame $t + 1$ can include all successful requests from these forced arrival slots and can still be communicated to the stations on time, i.e. before frame $t + 1$ actually starts. Specifically, this ensures that arrivals during these forced arrival slots of frame t can potentially depart in frame $t + 1$.

This implies though that, in case of the flexible boundary model, we must be careful about the exact location of the additional arrival slots. If they are located early within a frame, they may still be included in a schedule for the next frame. If however, such an additional arrival slot is located at the end of frame t, the corresponding request cannot be included in a schedule for frame $t + 1$ and must await the schedule for frame $t + 2$. In this paper, we have taken an optimistic viewpoint and have assumed that all successful reservations in the additional arrival slots of frame t can already be scheduled in frame $t+1$. More realistic models which account for the exact location of these additional arrival slots are more involved and are part of further research, as discussed in Section 6.

It remains to comment on the independence of the Y_{ti}, as assumed in both models. Clearly, the correlation between the Y_{ti} depends on the exact way in which the request procedure is organised. For cable networks, the requests are usually transmitted in contention with other stations and based on ALOHA or contention trees. These procedures have a considerable randomness in the order in which stations are actually successful: the time that a station has already been active in the contention procedure is not very significant as to its chances of being the next station to successfully transmit its request. This implies that the independence assumption made should be a good approximation to reality, see e.g. Boxma et al. [3].

In Section 4 we will give an exact analysis of the delay properties of both the fixed and flexible boundary model. In order to do so, we must first characterise the backlog in either model, which is the topic of the next section.

3 Backlog

3.1 Backlog in fixed boundary model

Let us denote by Y a random variable having the same distribution as the number of arriving packets during one arrival slot (i.e. $Y_{ti} \stackrel{d}{=} Y$ for all t and i where $\stackrel{d}{=}$ denotes equality in distribution), and we denote by $Y(z)$ the corresponding probability generating function (pgf)

$$Y(z) = \sum_{k=0}^{\infty} \Pr[Y = k] z^k, \quad |z| \leq 1.$$

Clearly, to have stability, it is required that the number of arriving packets is less than the maximum number of packets that can be transmitted, and hence Y should satisfy

$$c\mathbb{E}Y < s. \tag{3}$$

We have denoted the backlog at the beginning of frame t by X_t. Then $\{X_t, t \in \mathbb{Z}^+\}$ constitutes a discrete-time Markov chain, with transitions governed by (1). As is easily verified, the following conditional expectation holds

$$\mathbb{E}(z^{X_{t+1}} | X_t = k) = \begin{cases} Y(z)^c, & k < s, \\ z^{k-s} Y(z)^c, & k \geq s. \end{cases} \tag{4}$$

For reasons of brevity, we introduce the random variable A denoting the c-fold convolution of Y, that is, the pgf of A is given by $A(z) = Y(z)^c$.

Let X denote the steady-state distribution of the backlog, with

$$x_k = \Pr[X = k] = \lim_{t \to \infty} \Pr[X_t = k], \quad k = 0, 1, 2, \ldots$$

From (4) it follows that the pgf of X is given by (see e.g. [4])

$$X(z) = \frac{A(z) \sum_{k=0}^{s-1} [z^s - z^k] x_k}{z^s - A(z)}, \quad |z| \leq 1. \tag{5}$$

In this expression there are still s unknowns $x_0, \ldots, x_{s-1}$, which can be found using the following classical approach (see e.g. Bailey [2]). With Rouché's theorem, it can be shown that the denominator of (5) has s zeros on or within the unit circle $|z| \leq 1$. Since a pgf is analytic and well-defined in $|z| \leq 1$, the numerator of $X(z)$ should vanish at each of the zeros. This gives s equations. One of the zeros equals 1, and leads to a trivial equation. However, the normalization condition $X(1) = 1$ provides an additional equation. Using l'Hôpital's rule, this condition is found to be

$$s - \mathbb{E}A = \sum_{k=0}^{s-1} x_k(s-k), \tag{6}$$

which equates two expressions for the mean number of unused departure slots per frame.

Explicit expressions for the moments of the backlog can be obtained by taking derivatives of $X(z)$. For example, evaluating the first derivative of $X(z)$ at $z = 1$ yields

$$\mathbb{E}X = \frac{\mathbb{V}\text{ar}A}{2(s - \mathbb{E}A)} + \frac{s + \mathbb{E}A}{2} - \sum_{k=0}^{s-1} \frac{x_k(s-k)^2}{2(s - \mathbb{E}A)}. \tag{7}$$

So far we looked at the backlog at the beginning of a frame. We can also model the behavior of the backlog throughout a frame. Denote by $X_{[n]}$, $n = 1, 2, \ldots, f$, the steady-state backlog at the end of the n-th slot of a frame. The first c slots of a frame are arrival slots. This implies that the pgf of $X_{[n]}$ is given by

$$X_{[n]}(z) = X(z)Y(z)^n, \quad n = 1, \ldots, c. \tag{8}$$

The remaining s slots are departure slots, yielding

$$\mathbb{E}[z^{X_{[c+n]}} | X = k] = \begin{cases} A(z), & k < n, \\ A(z)z^{k-n}, & k \geq n. \end{cases} \tag{9}$$

Summing over all possible values of X then gives

$$X_{[c+n]}(z) = A(z)\Big[\sum_{k=0}^{n-1} x_k + \frac{1}{z^n}[X(z) - \sum_{k=0}^{n-1} x_k z^k]\Big], \quad n = 1, \ldots, s. \tag{10}$$

The expectation of the steady-state backlog throughout a frame then follows from evaluating the first derivative of (8) and (10) at $z = 1$. That is

$$\mathbb{E}X_{[n]} = \begin{cases} \mathbb{E}X + n\mathbb{E}Y, & n = 1, \ldots, c, \\ \mathbb{E}X + \mathbb{E}A - n + c + \sum_{k=0}^{n-c-1} x_k(n-c-k), & n = c+1, \ldots, f. \end{cases} \tag{11}$$

3.2 Backlog in flexible boundary model

As for the fixed boundary model, the first c slots of a frame are arrival slots. However, the unused departure slots can be turned into arrival slots as well. We have referred to this procedure as the *flexible boundary model*. The extra arrival slots are scheduled at the end of a frame. So, within a frame, the c forced arrival slots are scheduled first, then the departure slots (if any), and finally the additional arrival slots (if any). Stability condition (3) still holds and is equivalent to requiring c to be smaller than $f/(\mathbb{E}Y + 1)$. Although from a practical perspective it is more natural to schedule the additional arrival slots at the beginning of a frame, we choose this type of scheduling to simplify the analysis of the packet delay later on.

With $\tilde{X}_t$ representing the backlog at the beginning of frame t, $\{\tilde{X}_t, t \in \mathbb{Z}^+\}$ constitutes a discrete-time Markov chain, with transitions governed by (2). Note that the following conditional expectation holds

$$\mathbb{E}(z^{\tilde{X}_{t+1}} | \tilde{X}_t = k) = \begin{cases} Y(z)^{f-k}, & k < s, \\ z^{k-s} A(z), & k \geq s. \end{cases} \tag{12}$$

Because in the flexible boundary model all slots are used, the mean number of arrival slots per frame, denoted by c^*, is fixed and independent of c, i.e.

$$c^* = \frac{f}{\mathbb{E}Y + 1}, \tag{13}$$

as each arrival slot requires $1 + \mathbb{E}Y$ slots in total: the arrival slot itself and $\mathbb{E}Y$ slots for transmitting the packets.

Let $\tilde{X}$ denote the steady-state distribution of the backlog, with

$$\tilde{x}_k = \Pr[\tilde{X} = k] = \lim_{t\to\infty} \Pr[\tilde{X}_t = k], \quad k = 0, 1, 2, \ldots.$$

From (12), it follows that the pgf of $\tilde{X}$ is given by

$$\tilde{X}(z) = \frac{A(z) \sum_{k=0}^{s-1} [z^s Y(z)^{s-k} - z^k] \tilde{x}_k}{z^s - A(z)}, \quad |z| \leq 1. \tag{14}$$

As in Section 3.1, the s zeros of $z^s - A(z)$ on or within the unit circle $|z| \leq 1$ can be used to determine $\tilde{x}_0, \ldots, \tilde{x}_{s-1}$. Using l'Hôpital's rule, the normalization condition $\tilde{X}(1) = 1$ reads

$$s - \mathbb{E}A = \sum_{k=0}^{s-1} \tilde{x}_k (s-k)(\mathbb{E}Y + 1), \tag{15}$$

which equates two expressions for the number of slots per frame that are used for arrivals and departures of packets that arrived in other than the c forced arrival slots.

The mean backlog in case of the flexible boundary model is given by

$$\begin{aligned} \mathbb{E}\tilde{X} &= \frac{\mathbb{V}\text{ar}A}{2(s - \mathbb{E}A)} + \frac{s + \mathbb{E}A}{2} - \sum_{k=0}^{s-1} \frac{\tilde{x}_k (s-k)^2 (1 + \mathbb{E}Y)}{2(s - \mathbb{E}A)} \\ &+ \frac{\mathbb{V}\text{ar}Y}{2(\mathbb{E}Y + 1)} + \mathbb{E}Y \Big[\sum_{k=0}^{s-1} \frac{\tilde{x}_k (s-k)^2 (1 + \mathbb{E}Y)}{2(s - \mathbb{E}A)} \Big]. \end{aligned} \tag{16}$$

Expression (16) consists of the mean backlog corresponding to the fixed boundary model (7) and non-negative terms that solely depend on the mean and variance of Y. This follows from the observation that

$$\sum_{k=0}^{s-1} \frac{x_k (s-k)^2}{2(s - \mathbb{E}A)} = \sum_{k=0}^{s-1} \frac{\tilde{x}_k (s-k)^2 (1 + \mathbb{E}Y)}{2(s - \mathbb{E}A)}, \tag{17}$$

as shown in Denteneer et al. [5].

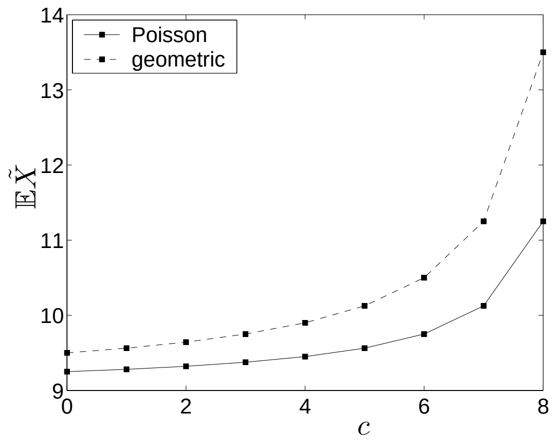

Figure 1: $\mathbb{E}\tilde{X}$ for $f = 18$, $\mathbb{E}Y = 1$ for Poisson and geometric distribution.

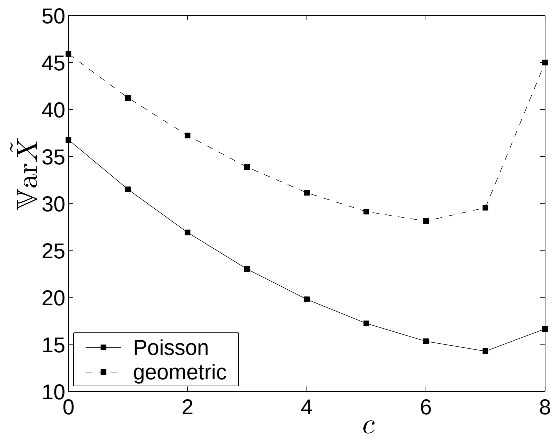

Figure 2: $\mathbb{V}\mathrm{ar}\tilde{X}$ for $f = 18$, $\mathbb{E}Y = 1$ for Poisson and geometric distribution.

Using the same notation as for the fixed boundary model, the behavior of the backlog throughout a frame follows from

$$\tilde{X}_{[n]}(z) = \tilde{X}(z)Y(z)^n, \quad n = 1, \ldots, c, \tag{18}$$

and

$$\mathbb{E}[z^{\tilde{X}_{[c+n]}}|\tilde{X} = k] = \begin{cases} Y(z)^{c+n-k}, & k < n, \\ A(z)z^{k-n}, & k \geq n, \end{cases} \tag{19}$$

and consequently,

$$\tilde{X}_{[c+n]}(z) = A(z)\Big[\sum_{k=0}^{n-1} \tilde{x}_k Y(z)^{n-k} + \frac{1}{z^n}[\tilde{X}(z) - \sum_{k=0}^{n-1} \tilde{x}_k z^k]\Big], \quad n = 1, \ldots, s. \tag{20}$$

Hence,

$$\mathbb{E}\tilde{X}_{[n]} = \begin{cases} \mathbb{E}\tilde{X} + n\mathbb{E}Y, & n = 1, \ldots, c, \\ \mathbb{E}Y[c + \sum_{k=0}^{n-c-1} \tilde{x}_k(n - c - k)] & \\ \quad +\mathbb{E}\tilde{X} - n + c + \sum_{k=0}^{n-c-1} \tilde{x}_k(n - c - k), & n = c + 1, \ldots, f. \end{cases} \tag{21}$$

Example 3.1. Consider a frame length of 18 slots, and Y distributed according to a Poisson and geometric distribution

$$\Pr[Y = k] = e^{-\lambda}\frac{\lambda^k}{k!}; \quad \Pr[Y = k] = (1 - p)p^k, \quad k = 0, 1, \ldots,$$

respectively, both with mean 1 ($\lambda = 1,\ p = 1/2$). The mean and variance of $\tilde{X}$ that correspond to these distributions are shown by Figure 1 and Figure 2 for various values of c. In terms of the mean backlog, having forced arrival slots at the beginning of the frame is disadvantageous. However, the variance of the backlog is reduced by increasing c. This stabilizing effect could be very welcome when the arrival process of packets is volatile.

Remark 3.2. In Jacquet et al. [8] a scheduling strategy called implicit framing is studied. For this strategy, no frame structure is used and priority is given to departure slots. Periods of consecutive arrival slots are implicitly closed by the first data packet to be transmitted. When all data packets have been transmitted, a new period of consecutive arrival slots, in which reservation takes place, is restarted. Note that such an implicit framing strategy is in fact the flexible boundary model with $f = 1$ and $c = 0$. Jacquet et al. [8] demonstrate that in terms of the average delay implicit framing is the best strategy within the class of the flexible boundary model. In the case of implicit framing, the pgf of $\tilde{X}$ reduces to

$$\tilde{X}(z) = \frac{\tilde{x}_0(zY(z) - 1)}{z - 1} = \frac{1 - zY(z)}{(1 - z)(\mathbb{E}Y + 1)}, \quad |z| \leq 1, \tag{22}$$

where $\tilde{x}_0$ equals $1/(\mathbb{E}Y + 1)$ according to the normalization condition (15). Note that $\tilde{X}$ can be interpreted as the residual lifetime of the random variable $Y + 1$. To see this, divide the time axis in cycles of one arrival slot plus the number of transmission slots Y granted during that arrival slot. The residual lifetime is an arbitrary point in a cycle, and since in every time slot during this cycle exactly one packet is transmitted, the residual lifetime equals the backlog. We stress though that implicit framing is less useful to model the data queue in cable networks, because the delay prevents that a request made by a station in time slot s is granted in time slot $s + 1$.

4 Packet delay

In deriving the packet delay distribution, the periodically scheduled departures cause some difficulties, as shown next. We first present some analysis that holds for both the fixed and flexible boundary model, after which we complete the analysis for both models separately.

Assume that the packets are transmitted in order of arrival. Tag an arbitrary packet, and let the random variable T denote the slot within the frame in which this packet arrives, $T \in \{1, 2, \ldots, f\}$. Assume that the packet arrives during slot $T = m$. Introduce $U_{[m]}$ as the number of packets present at the end of the frame that contribute to the tagged packet's delay. Then $U_{[m]}$ consists of the backlog at the end of the frame that was already present at the end of the previous frame, the packets that arrive in the same frame in arrival slots before T, and the packets that arrive within the same arrival slot but before the tagged packet. We then express $U_{[m]}$ in terms of two integer random variables $F_{[m]}$ and $R_{[m]}$

$$U_{[m]} = sF_{[m]} + R_{[m]}, \quad F_{[m]} \geq 0, \quad 0 \leq R_{[m]} \leq s - 1, \tag{23}$$

where $F_{[m]}$ denotes the number of complete frames enclosed in the tagged packet's delay, and $R_{[m]}$ the number of packets that will be transmitted during the same frame as the tagged packet, but before it. Introduce $D_{[m]}$ as the random variable representing the delay of a packet that arrives during arrival slot m, defined as

$$D_{[m]} = f - m + fF_{[m]} + c + R_{[m]} + 1. \tag{24}$$

That is, $f - m$ slots till the beginning of the next frame, $F_{[m]}$ frames, c forced arrival slots, $R_{[m]}$ slots within the frame of transmission, and the actual transmission slot of the tagged packet. The pgf of $D_{[m]}$ then reads

$$\begin{aligned} D_{[m]}(z) &= \sum_{i=0}^{\infty} \Pr[D_{[m]} = i] z^i \\ &= z^{f-m+c+1} \sum_{j=0}^{\infty} \sum_{k=0}^{s-1} \Pr[F_{[m]} = j, R_{[m]} = k] z^{fj+k} \\ &= z^{f-m+c+1} \sum_{j=0}^{\infty} \sum_{k=0}^{s-1} \Pr[U_{[m]} = sj + k] z^{fj+k}, \quad |z| \leq 1. \end{aligned} \tag{25}$$

From (25) it follows that

$$D_{[m]}(z^s) = z^{s(f-m+c+1)} \sum_{k=0}^{s-1} z^{sk} \vartheta_{mk}(z), \quad |z| \leq 1, \tag{26}$$

where the functions $\vartheta_{mk}(z)$ are defined as

$$\vartheta_{mk}(z) = \sum_{j=0}^{\infty} z^{sfj} \Pr(U_{[m]} = sj + k). \tag{27}$$

The problem now is that (27) does not allow a direct substitution of the pgf of $U_{[m]}$. To circumvent this, we use a basic approach that can be found in e.g. Bruneel & Kim [4] or Kang & Steyaert [12].

4.1 Basic approach

Substituting $l = sj + k$ in (27) yields

$$\vartheta_{mk}(z) = \sum_{l=0}^{\infty} z^{(l-k)f} \Pr(U_{[m]} = l) \sum_{j=-\infty}^{\infty} \delta(l - sj - k), \tag{28}$$

with $\delta(n)$ the Kronecker delta function, which equals 1 for $n = 0$ and 0 for all other n. Now invoke the following property

Property 4.1.

$$\frac{1}{s} \sum_{t=0}^{s-1} a^{tk} = \sum_{j=-\infty}^{\infty} \delta(k - js),$$

where $a = \exp(2\pi i/s)$, *i the imaginary unit, and k and s integer values. The sum on the left-hand side is zero unless k is a multiple of s.* □

Using Property 4.1 we obtain

$$\begin{aligned} \vartheta_{mk}(z) &= \sum_{l=0}^{\infty} z^{(l-k)f} \Pr(U_{[m]} = l) \frac{1}{s} \sum_{t=0}^{s-1} a^{t(l-k)} \\ &= \frac{z^{-kf}}{s} \sum_{t=0}^{s-1} a^{-tk} \sum_{l=0}^{\infty} \Pr(U_{[m]} = l) z^{fl} a^{tl} \\ &= \frac{z^{-kf}}{s} \sum_{t=0}^{s-1} a^{-tk} U_{[m]}(a^t z^f). \end{aligned} \tag{29}$$

Substituting (29) into (26) yields

$$
\begin{aligned}
D_{[m]}(z^s) &= z^{s(f-m+c+1)} \sum_{k=0}^{s-1} z^{sk} \frac{z^{-kf}}{s} \sum_{t=0}^{s-1} a^{-tk} U_{[m]}(a^t z^f) \\
&= \frac{z^{s(f-m+c+1)}}{s} \sum_{t=0}^{s-1} U_{[m]}(a^t z^f) \sum_{k=0}^{s-1} (z^{-c} a^{-t})^k \\
&= \frac{z^{s(f-m+c+1)}}{s} \sum_{t=0}^{s-1} U_{[m]}(a^t z^f) \frac{1-(z^{-c}a^{-t})^s}{1-z^{-c}a^{-t}}, \quad |z| \le 1. \qquad (30)
\end{aligned}
$$

Expression (30) gives an explicit formula for the pgf of the packet delay once the pgf of $U_{[m]}$ is known. This leaves us to specify the latter, for which we give separate derivations for the fixed and flexible boundary model.

4.2 Packet delay in fixed boundary model

Let D denote the packet delay for an arbitrary packet. Let Z_0 denote the backlog at the end of a frame that was already present the frame before, and Z_1 the number of packets within the tagged packet's arrival slot arriving before it. The pgf of Z_0 and Z_1 are given by

$$
Z_0(z) = \frac{1}{z^s}\left(X(z) + \sum_{k=0}^{s-1} x_k [z^s - z^k]\right), \qquad (31)
$$

and

$$
Z_1(z) = \frac{1-Y(z)}{(1-z)\mathbb{E}Y}. \qquad (32)
$$

The pgf of $U_{[m]}$ is then simply given by

$$
U_{[m]}(z) = Z_0(z) Y(z)^{m-1} Z_1(z), \quad m = 1, \ldots, c. \qquad (33)
$$

Combining (30), (33) and $\Pr[T = m] = 1/c$ for $m = 1, \ldots, c$, yields the following explicit expression for the pgf of the packet delay

$$
\begin{aligned}
D(z^s) &= \frac{1}{c} \sum_{m=1}^{c} D_{[m]}(z^s), \\
&= \frac{1}{sc} \sum_{t=0}^{s-1} \frac{1-(a^t z^c)^{-s}}{1-(a^t z^c)^{-1}} \left\{ z^{s(f+1)} Z_0(a^t z^f) Z_1(a^t z^f) \frac{z^{sc} - A(a^t z^f)}{z^s - Y(a^t z^f)} \right\}, \quad |z| \le 1. \qquad (34)
\end{aligned}
$$

The mean packet delay follows from

$$
\mathbb{E}D = \frac{1}{s} \frac{d}{dz} D(z^s) \bigg|_{z=1}, \qquad (35)
$$

which yields after tedious but straightforward calculations

$$
\mathbb{E}D = f + \frac{f \mathbb{V}\text{ar}A}{2\mathbb{E}A(s-\mathbb{E}A)} + \frac{1+\mathbb{E}Y}{\mathbb{E}Y} \left[\frac{s}{2} - \sum_{k=0}^{s-1} \frac{x_k (s-k)^2}{2(s-\mathbb{E}A)} \right]. \qquad (36)
$$

Remark 4.2. The mean delay can be alternatively derived using Little's law. The backlog at the beginning of an arbitrary slot is given by

$$\frac{1}{f}\sum_{n=1}^{f}\mathbb{E}X_{[n]}, \tag{37}$$

where $\mathbb{E}X_{[n]}$ as given by (11). The average arrival rate of packets per slot equals $c\mathbb{E}Y/f$. Dividing (37) through this rate then yields (36).

4.3 Packet delay in flexible boundary model

For the flexible boundary model, the derivation of $\tilde{U}_{[m]}(z)$ is somewhat more involved, since all slots within a frame are potential arrival slots. We first consider the case that $c \geq 1$, while $c = 0$ is covered at the end of this section. Distinguish two events: (a) the tagged packet arrives in one of the forced arrival slots, and (b) the tagged packet arrives in one of the additional arrival slots. Event (a) provides us no extra information about the backlog at the beginning of a frame, since the c forced arrival slots are scheduled every frame. Thus

$$\tilde{U}_{[m]}(z) = \tilde{Z}_0(z)Y(z)^{m-1}Z_1(z), \quad m = 1, \ldots, c, \tag{38}$$

where

$$\tilde{Z}_0(z) = \frac{1}{z^s}(\tilde{X}(z) + \sum_{k=0}^{s-1}\tilde{x}_k[z^s - z^k]). \tag{39}$$

Event (b) does provide extra information about the backlog at the beginning of the frame. We know that Z_0 equals zero, otherwise there would be no extra arrival slots. Further, consider the case that the tagged packet arrives in slot $c+1$. This implies that $\tilde{X}$ equals zero. Hence, $\tilde{U}_{[c+1]}$ consists of A and Z_1. Now consider the packet arriving in slot $c+2$. This implies that $\tilde{X}$ equals either zero or one. In the first case it holds that $\tilde{U}_{[c+2]} = A + Y + Z_1$, and in the latter case $\tilde{U}_{[c+2]} = A + Z_1$. Similar reasoning leads to the following expression

$$\tilde{U}_{[m]}(z) = A(z)Z_1(z)\frac{\sum_{k=0}^{m-c-1}\tilde{x}_k Y(z)^{m-c-1-k}}{\sum_{k=0}^{m-c-1}\tilde{x}_k}, \quad m = c+1, \ldots, f. \tag{40}$$

Finally, the distribution of T can be determined as follows. Remember that the extra arrival slots are scheduled at the end of a frame. If a packet arrives in slot m, $m \in \{c+1, \ldots, f\}$ of a frame, this particular frame has at least $f - m + c + 1$ arrival slots, and thus at most a backlog of $m - c - 1$ packets at the beginning of the frame. This gives

$$\Pr[T = m] = \begin{cases} \frac{1}{c^*}, & m = 1, \ldots, c, \\ \frac{1}{c^*}\sum_{k=0}^{m-c-1}\tilde{x}_k, & m = c+1, \ldots, f. \end{cases} \tag{41}$$

Combining (30), (38) and (40), and conditioning on the arrival slot distribution given by (41) yields an explicit expression for the pgf of the packet delay

$$
\begin{aligned}
\tilde{D}(z^s) &= \sum_{m=1}^{f} \Pr[T=m]\tilde{D}_{[m]}(z^s), \\
&= \sum_{m=1}^{c} \frac{1}{c^*}\tilde{D}_{[m]}(z^s) + \sum_{m=c+1}^{f} \frac{1}{c^*} \sum_{k=0}^{m-c-1} \tilde{x}_k \tilde{D}_{[m]}(z^s) \\
&= \frac{1}{sc^*} \sum_{t=0}^{s-1} \frac{1-(a^t z^c)^{-s}}{1-(a^t z^c)^{-1}} \Big\{ z^{s(f+1)} \tilde{Z}_0(a^t z^f) Z_1(a^t z^f) \frac{z^{sc} - A(a^t z^f)}{z^s - Y(a^t z^f)} \\
&+ z^{s(c+1)} A(a^t z^f) Z_1(a^t z^f) \frac{\sum_{k=0}^{s-1} \tilde{x}_k [z^{s(s-k)} - Y(a^t z^f)^k]}{z^s - Y(a^t z^f)} \Big\}, \quad |z| \le 1. \qquad (42)
\end{aligned}
$$

From (35), it follows that

$$
\mathbb{E}\tilde{D} = \frac{\mathbb{E}Y + 1}{\mathbb{E}Y} \Big\{ \mathbb{E}\tilde{X} + \frac{(s+1)\mathbb{E}A - s^2}{2f} + \sum_{k=0}^{s-1} \frac{\tilde{x}_k (s-k)^2 (1+\mathbb{E}Y)}{2f} \Big\}. \qquad (43)
$$

Again note that (43) can be alternatively determined by applying Little's law.

In case $c = 0$, the basic approach as described in Section 4.1 is not needed. It is then straightforward to derive that

$$
\tilde{U}_{[m]}(z) = Z_1(z) \frac{\sum_{k=0}^{m-1} x_k Y(z)^{m-1-k}}{\sum_{k=0}^{m-1} x_k}; \quad \Pr[\tilde{T} = m] = \frac{1}{c^*} \sum_{k=0}^{m-1} x_k,
$$

and

$$
\tilde{D}(z) = \frac{Z_1(z)}{f c^*} \sum_{m=1}^{f} z^{f-m+1} \sum_{k=0}^{m-1} x_k Y(z)^{m-1-k}. \qquad (44)
$$

Remark 4.3. We have derived the pgf of the packet delay for the fixed and flexible boundary model in (34) and (42), respectively. To find the underlying packet delay distribution we use a technique of Abate and Whitt [1]. A distribution $\{p_k\}$ can be retrieved from its pgf $P(z)$ via

$$
p_k = \frac{1}{2\pi i} \oint_{C_r} \frac{P(z)}{z^{k+1}} dz, \qquad (45)
$$

where C_r is a circle about the origin of radius $r,\ 0 < r < 1$. Abate and Whitt [1] approximate (45) using the trapezoidal rule with a step size of π/k as

$$
\hat{p}_k = \frac{1}{2\pi r^k} \sum_{j=1}^{2k} (-1)^j \mathrm{Re}(P(re^{ij\pi/k})), \qquad (46)
$$

and derive for $0 < r < 1,\ k \ge 1$ the following error bound

$$
|p_k - \hat{p}_k| \le \frac{r^{2k}}{1 - r^{2k}}. \qquad (47)
$$

For practical purposes one can think of the error bound as r^{2k}, because $r^{2k}/(1-r^{2k}) \approx r^{2k}$ for r^{2k} small. To have accuracy up to the γth decimal, we let $r = 10^{-\gamma/2k}$. In the upcoming numerical examples, we set γ equal to 7.

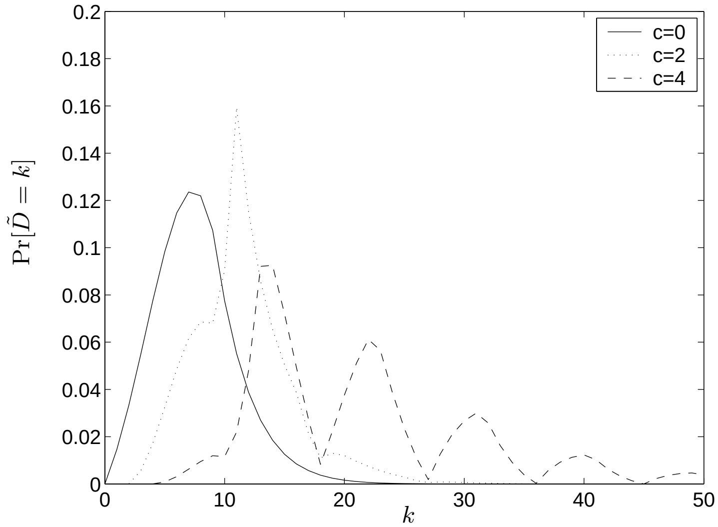

Figure 3: Distribution $\Pr[\tilde{D}=k]$, $f=9$, $c=0,2,4$, Y geometrically distributed with mean 1.

Example 4.4. The distribution of the packet delay has a characteristic form. For $f=9$, Y geometrically distributed with mean one, Figure 3 displays the packet delay distribution for $c=0,2,4$, where we have used the method described in Remark 4.3. First note that the minimum delay corresponds to a packet that arrives in the last slot of a frame and is immediately transmitted in the slot $c+1$ of the next frame. The oscillating effect is due to the frame structure, and becomes stronger for higher values of c.

5 Numerical results

In this section we first present a numerical comparison between the fixed and flexible boundary model. Next, we investigate the impact of different values of c for the flexible boundary model on various backlog and delay characteristics.

5.1 Fixed versus flexible boundary model

We assume that the load, defined as the mean number of packets arriving per frame, is the same for the fixed and flexible boundary model, being $c\mathbb{E}Y$ and $c^*\mathbb{E}Y = f\mathbb{E}Y/((1+\mathbb{E}Y))$, respectively. Thus, for a fair comparison, we choose the appropriate values of $\mathbb{E}Y$ for which the load is the same for both models. For convenience, we further assume that Y is Poisson distributed.

Figure 4 and Figure 5 display the mean and variance of the packet delay for $f=9$, $c=2$ and various load values. For a load of 2, $\mathbb{E}Y=1$ for the fixed boundary model and $\mathbb{E}Y=2/7$ ($c^*=7$, $c^*\mathbb{E}Y=2$) for the flexible boundary model. In terms of the mean packet delay, the flexible boundary model clearly outperforms its fixed counterpart.

For the flexible boundary model, a low load yields relatively many unused departure slots that are used as additional arrival slots. The variation in arrival slots per frame inherent to such type of scheduling then causes a higher packet delay variance than in case of the fixed boundary model. When the load gets higher, this variation in arrival slots can give relief to the system, leading to a smaller overall variance.

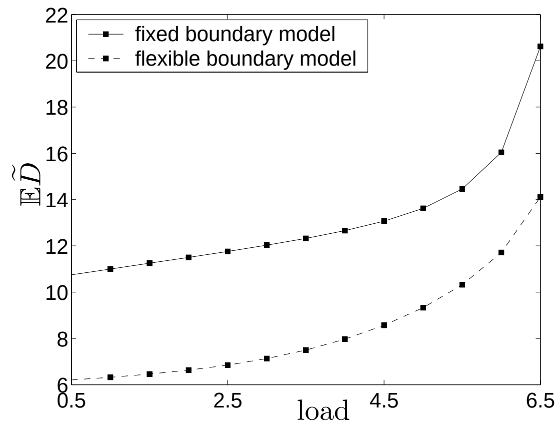

Figure 4: Mean packet delay, fixed vs. flexible boundary model, $f = 9$, $c = 2$, Y Poisson distributed.

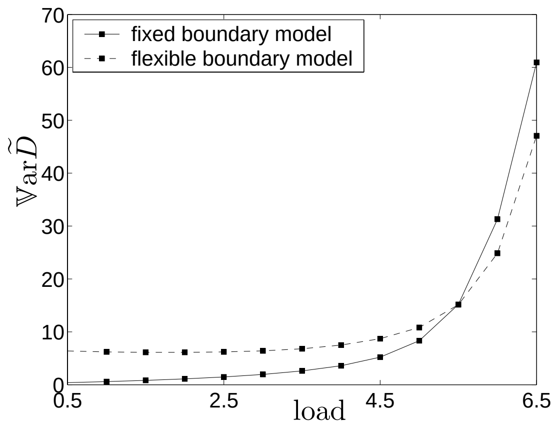

Figure 5: Packet delay variance, fixed vs. flexible boundary model, $f = 9$, $c = 2$, Y Poisson distributed.

5.2 Influence of c in flexible boundary model

We now investigate the impact of different values of c for the flexible boundary model on various backlog and delay characteristics. Table 1 contains backlog characteristics for $f = 9$, $c = 0, 2, 4$, and Y is Poisson and geometrically distributed with mean 1. Table 2 contains delay characteristics for the same settings.

		$\mathbb{E}\tilde{X}$	$\mathbb{V}\mathrm{ar}\tilde{X}$	$\Pr[\tilde{X} > 10]$	$\Pr[\tilde{X} > 20]$	$\Pr[\tilde{X} > 50]$
Poisson	$c = 0$	4.75	11.75	.0639	.0003	.0000
	$c = 2$	4.95	7.97	.0408	.0001	.0000
	$c = 4$	6.75	10.93	.1245	.0019	.0002
geometric	$c = 0$	5.00	16.67	.1042	.0026	.0001
	$c = 2$	5.40	14.07	.0995	.0020	.0000
	$c = 4$	9.00	34.63	.3197	.0471	.0064

Table 1: Characteristics of the backlog for $f = 9$ and $\mathbb{E}Y = 1$.

		$\mathbb{E}\tilde{D}$	$\mathbb{V}\mathrm{ar}\tilde{D}$	$\Pr[\tilde{D} > 10]$	$\Pr[\tilde{D} > 20]$	$\Pr[\tilde{D} > 30]$
Poisson	$c = 0$	6.92	7.60	.0926	.0020	.0000
	$c = 2$	8.57	8.82	.5437	.0039	.0001
	$c = 4$	13.66	32.03	.9550	.2800	.0327
geometric	$c = 0$	7.63	11.84	.1767	.0028	.0003
	$c = 2$	11.46	17.10	.6075	.0353	.0014
	$c = 4$	21.40	96.86	.9568	.4855	.1812

Table 2: Characteristics of the packet delay for $f = 9$ and $\mathbb{E}Y = 1$.

As we have seen in Example 3.1, increasing c is disadvantageous in terms of the mean backlog, while the variance of the backlog is oftentimes reduced due to the stabilizing effect on the arrival process. The same can be seen from the results in Table 1. Increasing c reduces the flexibility of the system, which is the reason for both the mean and variance of the backlog to increase when c gets large, $c = 4$ in this example.

The mean and variance give only partial information on the underlying distribution function. We therefore consider some excess probabilities. Note that in Table 1, the probability that $\tilde{X}$ gets larger than 10 is the smallest for $c = 2$, for both the Poisson and geometric distribution. Depending on which performance characteristic one is interested in, one can determine the optimal value of c.

For the delay characteristics in Table 2 we do not see a stabilizing effect. This is mainly due to our definition of delay that includes the delay of the arrival slot till the beginning of the next frame. In this way small values of c are favored, since small values of c imply many additional arrival slots that bring along a smaller delay till the beginning of the next frame.

6 Further research

We have mentioned earlier that, in case of the flexible boundary model, one must be careful about the exact location of the additional arrival slots. The approach in this paper sketches an optimistic scenario, since the packets that arrive in the additional arrival slots can all be transmitted at the beginning of the next frame. A pessimistic scenario would be to assume that the packets arriving during the additional arrival slots can only be transmitted from the beginning of the second next frame. The recursion describing the backlog then becomes

$$\begin{pmatrix} X_{t+1} \\ R_{t+1} \end{pmatrix} = \begin{pmatrix} (X_t - s)^+ + \sum_{i=1}^{c} Y_{ti}^{(1)} + R_t \\ \sum_{i=1}^{(s-X_t)^+} Y_{ti}^{(2)} \end{pmatrix}, \tag{48}$$

where again

$$Y_{ti}^{(j)} \text{i.i.d.}, \quad Y_{ti}^{(j)} \stackrel{d}{=} Y, \quad j = 1, 2, \tag{49}$$

X_t the part of the backlog at the beginning of frame t that can be transmitted in frame t, and R_t the part of the backlog at the beginning of frame t that cannot be transmitted in frame t.

It is possible though that a scheduler could locate these extra request slots somewhat more efficiently. This depends on the exact location of the last slot within a frame that can carry a request that can be scheduled in the next frame, slot k say. Assuming that all arrival slots are scheduled at the beginning of a frame, the backlog can be described via the recursion

$$\begin{pmatrix} X_{t+1} \\ R_{t+1} \end{pmatrix} = \begin{pmatrix} (X_t - s)^+ + \sum_{i=1}^{c} Y_{ti}^{(1)} + \sum_{i=1}^{\min\{k,(s-X_t)^+\}} Y_{ti}^{(2)} + R_t \\ \sum_{i=1}^{\max\{0,(s-X_t)^+-k\}} Y_{ti}^{(3)} \end{pmatrix}, \tag{50}$$

where again

$$Y_{ti}^{(j)} \text{i.i.d.}, \quad Y_{ti}^{(j)} \stackrel{d}{=} Y, \quad j = 1, 2, 3. \tag{51}$$

The analysis of equations (48) and (50) is an interesting subject for further research.

Acknowledgement

The authors like to thank Bart Steyaert for helpful comments.

References

[1] Abate, J., W. Whitt (1992). Numerical inversion of probability generating functions, *Operations Research Letters* **12**: 245-251.

[2] Bailey, N.T.J. (1954). On queueing processes with bulk service, *Journal of the Royal Statistical Society* **16**: 80-87.

[3] Boxma, O., D. Denteneer, J. Resing (2003). Delay models for contention trees in closed populations, *Performance Evaluation* **53**: 169-185.

[4] Bruneel, H., B.G. Kim (1993). *Discrete-Time Models for Communication Systems including ATM*, Kluwer Academic Publishers, Dordrecht.

[5] Denteneer, D., J. van Leeuwaarden, J. Resing (2003). Bounds for a discrete-time multi server queue with an application to cable networks, In *Proceedings of ITC 18*, Berlin, p. 601-612.

[6] Golmie, N., S. Masson, G. Pieris, D.H. Su (1997). A MAC protocol for HFC networks: Design issues and performance evaluation. *Computer Communications* **20**: 1042-1050.

[7] Golmie, N., Y. Santillan, D.H. Su (1999). A review of contention resolution algorithms for the IEEE 802.14 networks. *IEEE Communication Surveys*: 2-12.

[8] Jacquet, P., P. Mühlethaler, P. Robert (2001). Framing protocols on upstream channel in CATV networks: asymptotic average delay analysis, *INRIA Report 4114*, Rocquencourt.

[9] Norimatsu, T, H. Takagi, H.R. Gail (2002). Performance analysis of the IEEE 1394 serial bus. *Performance Evaluation* **50**: 1-26.

[10] Resing, J.A.C., L. Örmeci (2003). A tandem queueing model with coupled processors, *Operations Research Letters* **31**(5): 383-389.

[11] Sala, D., J.O. Limb, S.U. Khaunte (1998). Adaptive control mechanism for cable modem MAC protocols, In *Proceedings of IEEE Infocom 1998*, San Francisco, p. 1392-1399.

[12] Kang, K., B. Steyaert (1999). Bounds analysis for WRR scheduling in a statistical multiplexer with bursty sources. *Telecommunications Systems* **12**: 123-147.

[13] Van Eenige, M.J.A. (1996). *Queueing Systems with Periodic Service*, Ph.D. thesis, Technical University Eindhoven, Eindhoven.

[14] Zhao, Y.Q. and L.L. Campbell (1996). Equilibrium probability calculations for a discrete-time bulk queue model. *Queueing Systems* **22**: 189-198.