

# A discrete-time priority queue with train arrivals

Joris Walraevens, Sabine Wittevrongel and Herwig Bruneel

SMACS Research Group

Department of Telecommunications and Information Processing (IR07)

Ghent University - UGent

Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium.

Phone: +32-9-2648902, Fax: +32-9-2644295

E-mail: {jw,sw,hb}@telin.UGent.be

## Abstract

We analyze a discrete-time priority queue with train arrivals. Messages of a variable number of fixed-length packets belonging to two classes arrive to the queue at the rate of one packet per slot. We assume geometrically distributed message lengths. Packets of the first class have transmission priority over the packets of the other class. By using probability generating functions, some performance measures such as the moments of the packet delay are calculated. The impact of the priority scheduling discipline and the correlation in the arrival process is shown by some numerical examples.

**Keywords:** discrete-time priority queue, correlated arrivals, probability generating functions

## 1 Introduction

In recent years, there has been much research on incorporating multimedia applications in packet-based networks (e.g. IP networks). Different types of traffic need different QoS (Quality-of-Service) standards. E.g. for real-time (interactive) applications (telephony, multimedia, gaming, ...), it is important that the mean delay and delay jitter are minimal, while this is of minor importance for data applications (file transfers, E-mail, ...). In order to guarantee acceptable delay boundaries to real-time traffic, several scheduling schemes – for switches, routers, ... – have been proposed and analyzed, each one with its own specific algorithmic and computational complexity. The most drastic one is *HOL* (Head-Of-the-Line) *priority*. With this scheduling discipline, as long as real-time (or

high-priority) packets are present in the system, they are transmitted. Data (or low-priority) packets can thus only be transmitted when no high-priority traffic is present.

In the related literature, there have been a number of contributions with respect to the analysis of HOL priority queues. An overview of some basic HOL priority queueing models can be found in [5, 9] and the references therein. Discrete-time HOL priority queues with deterministic service times equal to one slot have been studied in [6, 7, 10, 13]. In [6], the steady-state system content and delay in the case of a multiserver queue are studied. Mehmet Ali and Song [7] analyze the system content in a multiplexer with two-state on-off sources. The steady-state system content and the delay for Markov modulated high-priority interarrival times and geometrically distributed low-priority interarrival times are presented in [10]. Walraevens et al. [13] study the steady-state system content and packet delay, in the special case of an output queueing switch with Bernoulli arrivals.

In this paper, we assume an arrival process induced by a two-layered structure. Higher-layer *messages* consist of a number of lower-layer *packets*. We assume that the messages arrive to the system in the form of *trains* of packets. Since we will perform a discrete-time analysis, we assume time is divided into slots of equal length and the messages thus arrive to the queue at the rate of one packet per slot. The rationale for this arrival process is the layered structure of packet-based networks. Packets arriving to a queue are part of larger entities, the messages. These messages can e.g. be IP flows or IP sessions. From a lower layer, these messages are seen as *trains of packets*. E.g. IP flows are seen as trains of IP packets. Another example is IP-over-ATM, where the IP packets are in fact the messages in this example, while the ATM cells are the packets.

This two-layered structure introduces *time correlation* in the arrival process. Indeed, since the packets of a message arrive in consecutive slots, the number of packet arrivals in one slot depends on the number of arrivals in previous slots. FIFO queues with train arrivals are analyzed in [1, 2, 4, 14, 15]. Furthermore, priority queues with train arrivals and deterministic message lengths are studied in [3]. In the current paper, we analyze a discrete-time priority queue with train arrivals and geometrically distributed message lengths. The distributions of the message lengths may be class-dependent, i.e., the parameter of the geometric distribution may be different for the high-priority and low-priority message lengths. We will analyze the system contents (i.e., the number of packets in the system) as well as the packet delays (i.e., the number of slots a packet stays in the queue) of both the high-priority and low-priority class using *probability generating functions* (pgfs). This will eventually lead to explicit closed-form formulas for the means of these stochastic variables (and in most cases also for the higher

moments).

The remainder of this paper is structured as follows. In the next section, we present the mathematical model. In sections 3 and 4, we then analyze the steady-state system contents and packet delays of both classes. Some numerical examples are treated in section 5, while we conclude this paper in section 6.

## 2 Mathematical model

We consider a discrete-time single-server system with infinite buffer space. Time is assumed to be slotted. There are two types of messages arriving to the system, namely messages of class 1 and messages of class 2. The numbers of newly generated class- $j$  messages during consecutive slots are i.i.d. (independent and identically distributed). The numbers of newly generated class-1 and class-2 messages during slot  $k$  are denoted by  $b_{1,k}$  and  $b_{2,k}$  respectively. Their joint pgf is defined as

$$B(z_1, z_2) \triangleq \mathbb{E} \left[ z_1^{b_{1,k}} z_2^{b_{2,k}} \right].$$

Note that the numbers of generated messages of both classes during a slot may be correlated. The marginal pgfs of the number of newly generated messages of class  $j$  are denoted by  $B_j(z)$  ( $j = 1, 2$ ) and are given by  $B(z, 1)$  and  $B(1, z)$  respectively.

Each class- $j$  message is composed of a random number of packets which is assumed (shifted) geometrically distributed, i.e., its pgf is given by

$$L_j(z) = \frac{(1 - \sigma_j)z}{1 - \sigma_j z},$$

$j = 1, 2$ . The packets of a message arrive back to back at the rate of one packet per slot. The total numbers of class-1 and class-2 packets arriving during slot  $k$  are denoted by  $a_{1,k}$  and  $a_{2,k}$  respectively and their joint pgf is defined as

$$A_k(z_1, z_2) \triangleq \mathbb{E} \left[ z_1^{a_{1,k}} z_2^{a_{2,k}} \right].$$

The transmission times of the packets equal one slot and per slot one packet can be transmitted (if there is any).

Packets of class 1 have HOL priority over packets of class 2. Since each class- $j$  message consists

on average of  $L'_j(1) = 1/(1 - \sigma_j)$  packets, the load generated by class- $j$  packets equals

$$\rho_j = \frac{B'_j(1)}{1 - \sigma_j},$$

$j = 1, 2$ . We assume a stable system, i.e., the total load has to be smaller than 1:

$$\rho_T \triangleq \rho_1 + \rho_2 = \frac{B'_1(1)}{1 - \sigma_1} + \frac{B'_2(1)}{1 - \sigma_2} < 1. \quad (1)$$

### 3 System content

#### 3.1 Functional equation

In this section, we analyze the system content. We denote the system content of class-1 packets and class-2 packets at the beginning of slot  $k$  by  $u_{1,k}$  and  $u_{2,k}$  respectively. The following system equations are found:

$$\begin{aligned} u_{1,k+1} &= [u_{1,k} - 1]^+ + a_{1,k}; \\ u_{2,k+1} &= \begin{cases} [u_{2,k} - 1]^+ + a_{2,k} & \text{if } u_{1,k} = 0 \\ u_{2,k} + a_{2,k} & \text{if } u_{1,k} > 0, \end{cases} \end{aligned} \quad (2)$$

where  $[\cdot]^+$  denotes the maximum of the argument and 0. The first equation follows from the observation that the transmission of class-1 packets is not influenced by the class-2 packets. A class-2 packet on the other hand can only be transmitted during slot  $k$ , if there are no class-1 packets in the system at the beginning of this slot. This leads to the second equation.

Due to the correlation in the packet arrival process,  $(u_{1,k}, u_{2,k})$  does not form a Markovian state description of the system at the beginning of slot  $k$ . Therefore, we also include the numbers of packet arrivals of both classes during the previous slot into the description of the system. Because of the geometric distribution of the message lengths of both classes, this leads to the Markovian state description  $(a_{1,k-1}, u_{1,k}, a_{2,k-1}, u_{2,k})$  of the system. Besides (2), we then further need a relation between  $a_{j,k}$  and  $a_{j,k-1}$  for  $j = 1, 2$  to fully characterize the behavior of the Markov chain. We find

$$a_{j,k} = b_{j,k} + \sum_{i=1}^{a_{j,k-1}} c_{j,k}^{(i)}, \quad j = 1, 2. \quad (3)$$

The first term corresponds to the arrivals of first packets of newly generated messages during slot  $k$ ,

while the second term represents packet arrivals during slot  $k$  of messages which were already arriving during the previous slot. The  $c_{1,k}^{(i)}$ 's are i.i.d. random variables, which equal 1 or 0 with probabilities  $\sigma_1$  and  $1 - \sigma_1$  respectively. Likewise, the  $c_{2,k}^{(i)}$ 's are i.i.d. random variables, which equal 1 or 0 with probabilities  $\sigma_2$  and  $1 - \sigma_2$  respectively.

Equations (2) and (3) fully describe the behavior of the system. We introduce the following 4-dimensional pgf

$$P_k(x_1, z_1, x_2, z_2) \triangleq \mathbb{E} \left[ x_1^{a_{1,k-1}} z_1^{u_{1,k}} x_2^{a_{2,k-1}} z_2^{u_{2,k}} \right].$$

From equation (2) it follows that

$$\begin{aligned} P_{k+1}(x_1, z_1, x_2, z_2) = & \mathbb{E} \left[ (x_1 z_1)^{a_{1,k}} (x_2 z_2)^{a_{2,k}} z_2^{[u_{2,k-1}]^+} \mathbf{1}_{\mathbf{u}_{1,k}=\mathbf{0}} \right] \\ & + \mathbb{E} \left[ (x_1 z_1)^{a_{1,k}} z_1^{u_{1,k}-1} (x_2 z_2)^{a_{2,k}} z_2^{u_{2,k}} \mathbf{1}_{\mathbf{u}_{1,k}>\mathbf{0}} \right], \end{aligned}$$

with  $\mathbf{1}_X$  being the indicator function of  $X$ . Further, using expression (3) and explicitly taking into account that  $b_{1,k+1}$  and  $b_{2,k+1}$  are statistically independent of the other random variables and the fact that the  $c_{1,k}^{(i)}$ 's (the  $c_{2,k}^{(i)}$ 's respectively) are independent and Bernoulli distributed with parameter  $\sigma_1$  ( $\sigma_2$  respectively), this can be rewritten as

$$\begin{aligned} P_{k+1}(x_1, z_1, x_2, z_2) = & B(x_1 z_1, x_2 z_2) \left( \mathbb{E} \left[ \prod_{j=1}^2 (1 - \sigma_j + \sigma_j x_j z_j)^{a_{j,k-1}} z_2^{[u_{2,k-1}]^+} \mathbf{1}_{\mathbf{u}_{1,k}=\mathbf{0}} \right] \right. \\ & \left. + \mathbb{E} \left[ \prod_{j=1}^2 (1 - \sigma_j + \sigma_j x_j z_j)^{a_{j,k-1}} z_1^{u_{1,k}-1} z_2^{u_{2,k}} \mathbf{1}_{\mathbf{u}_{1,k}>\mathbf{0}} \right] \right). \end{aligned}$$

A little more algebra finally yields

$$\begin{aligned} P_{k+1}(x_1, z_1, x_2, z_2) = & \frac{B(x_1 z_1, x_2 z_2)}{z_1 z_2} [z_2 P_k(1 - \sigma_1 + \sigma_1 x_1 z_1, z_1, 1 - \sigma_2 + \sigma_2 x_2 z_2, z_2) \\ & + (z_1 - z_2) P_k(0, 0, 1 - \sigma_2 + \sigma_2 x_2 z_2, z_2) + z_1 (z_2 - 1) P_k(0, 0, 0, 0)], \end{aligned} \quad (4)$$

where we have used the property that  $a_{j,k-1} = 0$  if  $u_{j,k} = 0$  for  $j = 1, 2$ . In the steady state,  $P_k(x_1, z_1, x_2, z_2)$  and  $P_{k+1}(x_1, z_1, x_2, z_2)$  both converge to the same limiting function  $P(x_1, z_1, x_2, z_2)$ .

It then follows from (4) that this function must satisfy the following functional equation:

$$P(x_1, z_1, x_2, z_2) = \frac{B(x_1 z_1, x_2 z_2)}{z_1 z_2} [z_2 P(1 - \sigma_1 + \sigma_1 x_1 z_1, z_1, 1 - \sigma_2 + \sigma_2 x_2 z_2, z_2) + (z_1 - z_2) P(0, 0, 1 - \sigma_2 + \sigma_2 x_2 z_2, z_2) + z_1 (z_2 - 1) P(0, 0, 0, 0)]. \quad (5)$$

Equation (5) contains all information concerning the steady-state behavior of the system, although not in a transparent form. Nevertheless, several explicit results can be derived from it, which is the subject of the following subsections.

### 3.2 The arrival process revisited

Firstly, the joint pgf of the total numbers of packets of class 1 and class 2 arriving during a random slot in the steady state can be calculated. It is given by

$$\begin{aligned} A(z_1, z_2) &\triangleq \lim_{k \rightarrow \infty} A_k(z_1, z_2) = P(z_1, 1, z_2, 1) \\ &= B(z_1, z_2) P(1 - \sigma_1 + \sigma_1 z_1, 1, 1 - \sigma_2 + \sigma_2 z_2, 1) \\ &= B(z_1, z_2) A(1 - \sigma_1 + \sigma_1 z_1, 1 - \sigma_2 + \sigma_2 z_2) \end{aligned} \quad (6)$$

Successive applications of (6) lead to an explicit expression for  $A(z_1, z_2)$ , namely,

$$A(z_1, z_2) = \prod_{k=0}^{\infty} B(1 - \sigma_1^k + \sigma_1^k z_1, 1 - \sigma_2^k + \sigma_2^k z_2), \quad (7)$$

if the infinite product in the right-hand side converges. We will comment on this convergence later on. The marginal pgf of the number of class- $j$  packets arriving in a slot in the steady state follows from  $A(z_1, z_2)$  as

$$A_j(z) = \prod_{k=0}^{\infty} B_j(1 - \sigma_j^k + \sigma_j^k z), \quad (8)$$

$j = 1, 2$  and the marginal pgf of the total number of arriving packets in a slot equals

$$A_T(z) = A(z, z) = \prod_{k=0}^{\infty} B(1 - \sigma_1^k + \sigma_1^k z, 1 - \sigma_2^k + \sigma_2^k z).$$

We will now first briefly comment on the convergence of the infinite product in (7). Neuts [8]

has treated the convergence of this type of infinite product in case of a one-dimensional pgf, i.e., he investigated the convergence of the infinite product on the right-hand side of (8). The sole condition for convergence is the existence of the corresponding mean, or, in the context of equation (8), the finiteness of  $B'_j(1)$ . Extending the reasoning of [8] to two-dimensional pgfs, it can easily be found that a sufficient condition for convergence of the product in (7) is the finiteness of the mean total number of newly generated messages per slot.

From (8), the mean number of class- $j$  packet arrivals per slot can be found. It is given by

$$A'_j(1) = \frac{B'_j(1)}{1 - \sigma_j}, \quad (9)$$

while the mean total number of packet arrivals equals

$$A'_T(1) = \frac{B'_1(1)}{1 - \sigma_1} + \frac{B'_2(1)}{1 - \sigma_2}.$$

Since the mean total number of arrivals entering the buffer per slot must be strictly less than the maximal transmission rate (which equals one packet per slot for this model), we again find the stability condition (1). Other (cross-)moments of the packet arrival process can be obtained in a similar manner.

### 3.3 The system content

The main goal of this section is the analysis of the system content, which is now continued in this subsection. We first denote the joint pgf of the steady-state class-1 and class-2 system contents at the beginning of a random slot by  $U(z_1, z_2)$ . This pgf is given by

$$U(z_1, z_2) \triangleq \lim_{k \rightarrow \infty} \mathbb{E} [z_1^{u_{1,k}} z_2^{u_{2,k}}] = P(1, z_1, 1, z_2).$$

Expression (5) thus yields a functional equation for  $U(z_1, z_2)$ . In general, it does not seem to be possible though to derive an explicit expression for  $U(z_1, z_2)$  from this functional equation. By selecting only those  $(x_1, z_1, x_2, z_2)$ -values for which the  $P$ -functions on both sides of equation (5) have identical arguments, i.e., by choosing  $x_j = (1 - \sigma_j)/(1 - \sigma_j z_j) (= L_j(z_j)/z_j)$  for  $j = 1, 2$ , we can derive the following expression:

$$P\left(\frac{L_1(z_1)}{z_1}, z_1, \frac{L_2(z_2)}{z_2}, z_2\right) = \frac{B(L_1(z_1), L_2(z_2))}{z_2 [z_1 - B(L_1(z_1), L_2(z_2))]} \quad (10)$$

$$\times \left[ (z_1 - z_2)P\left(0, 0, \frac{L_2(z_2)}{z_2}, z_2\right) + z_1(z_2 - 1)P(0, 0, 0, 0) \right].$$

In the remainder of this subsection, we will firstly show that an (almost explicit) expression can be found for this function and secondly that this is sufficient to calculate the means of the system contents of both classes and of the total system content.

By applying Rouché's theorem, it can be proved that for a given value of  $z_2$  ( $|z_2| \leq 1$ ), the equation  $z_1 = B(L_1(z_1), L_2(z_2))$  has one solution in the unit circle for  $z_1$ , which we denote by  $Y(z_2)$ .  $Y(z)$  is thus implicitly defined as

$$Y(z) \triangleq B(L_1(Y(z)), L_2(z)), \quad |Y(z)| < 1 \text{ if } |z| < 1. \quad (11)$$

Since  $Y(z_2)$  is a zero of the denominator of the right-hand side of (10) and since a pgf remains finite for its arguments inside the complex unit circle,  $Y(z_2)$  must be a zero of the numerator as well. We thus find

$$P\left(0, 0, \frac{L_2(z_2)}{z_2}, z_2\right) = P(0, 0, 0, 0) \frac{Y(z_2)(z_2 - 1)}{z_2 - Y(z_2)}.$$

Substituting this result in equation (10) yields

$$P\left(\frac{L_1(z_1)}{z_1}, z_1, \frac{L_2(z_2)}{z_2}, z_2\right) = P(0, 0, 0, 0) \frac{B(L_1(z_1), L_2(z_2))(z_2 - 1)}{z_2 - Y(z_2)} \frac{z_1 - Y(z_2)}{z_1 - B(L_1(z_1), L_2(z_2))}. \quad (12)$$

$P(0, 0, 0, 0)$  can be derived by applying the normalization condition  $P(1, 1, 1, 1) = 1$ . Use of de l'Hôpital's rule gives the expected result for the probability of having an empty system:  $P(0, 0, 0, 0) = 1 - \rho_T$ . Before going into the calculation of the moments, we define the following three pgfs:

$$\begin{aligned} P_1(x, z) &\triangleq P(x, z, 1, 1), \\ P_2(x, z) &\triangleq P(1, 1, x, z), \\ P_T(x_1, x_2, z) &\triangleq P(x_1, z, x_2, z). \end{aligned} \quad (13)$$

$P_1(x, z)$  and  $P_2(x, z)$  are thus the steady-state joint pgfs of the number of packet arrivals of class 1 (class 2 respectively) per slot and the number of class-1 (class-2 respectively) packets stored in the buffer at the beginning of the following slot.  $P_T(x_1, x_2, z)$  is the steady-state joint pgf of the number of



packet arrivals of class-1 during a random slot, the number of packet arrivals of class-2 during that slot, and the total system content at the beginning of the following slot. In view of (12),  $P_1(L_1(z)/z, z)$ ,  $P_2(L_2(z)/z, z)$  and  $P_T(L_1(z)/z, L_2(z)/z, z)$  are given by

$$P_1\left(\frac{L_1(z)}{z}, z\right) = (1 - \rho_1) \frac{B_1(L_1(z))(z-1)}{z - B_1(L_1(z))}; \quad (14)$$

$$P_2\left(\frac{L_2(z)}{z}, z\right) = (1 - \rho_T) \frac{B_2(L_2(z))(z-1)}{z - Y(z)} \frac{Y(z) - 1}{B_2(L_2(z)) - 1}, \quad (15)$$

$$P_T\left(\frac{L_1(z)}{z}, \frac{L_2(z)}{z}, z\right) = (1 - \rho_T) \frac{B(L_1(z), L_2(z))(z-1)}{z - B(L_1(z), L_2(z))}. \quad (16)$$

Note that in order to obtain the expression of  $P_1(L_1(z)/z, z)$  from (12), de l'Hôpital's rule has to be applied. This calculation further needs the calculation of the first derivative of  $Y(z)$  in  $z = 1$ . From expression (11), we find

$$\begin{aligned} Y'(1) &= \frac{B'_1(1)}{1 - \sigma_1} Y'(1) + \frac{B'_2(1)}{1 - \sigma_2} \\ &= \frac{\rho_2}{1 - \rho_1}, \end{aligned} \quad (17)$$

since  $Y(1) = 1$  and in view of the definitions of the class-1 and class-2 loads. Note that  $Y(1) = 1$  since  $Y(z)$  is in fact a pgf (see [12] for a similar example).

Expressions (14), (15) and (16) will now be used for the calculation of the mean class-1 system content, the mean class-2 system content and the mean total system content respectively. The mean number of class-1 packets  $E[u_1]$  can be found as follows. First we take the first derivative with respect to  $z$  of both sides of (14) at  $z = 1$ . This leads to

$$\frac{\partial P_1}{\partial x}(1, 1) \frac{\sigma_1}{1 - \sigma_1} + \frac{\partial P_1}{\partial z}(1, 1) = \frac{2B'_1(1)(1 - B'_1(1)) + B''_1(1)}{2(1 - \sigma_1)(1 - \sigma_1 - B'_1(1))}.$$

The partial derivatives  $[\partial P_1/\partial x](1, 1)$  and  $[\partial P_1/\partial z](1, 1)$  equal  $A'_1(1)$  and  $E[u_1]$  respectively. Solving for  $E[u_1]$  and introducing  $\rho_1$ , we obtain the mean class-1 system content as

$$E[u_1] = \frac{\rho_1}{1 - \rho_1} - \frac{(1 - 2\sigma_1)\rho_1^2}{(1 - \sigma_1)(1 - \rho_1)} + \frac{B''_1(1)}{2(1 - \sigma_1)^2(1 - \rho_1)}. \quad (18)$$

The mean class-2 system content  $E[u_2]$  is calculated in a similar manner. Taking the first derivative

with respect to  $z$  of both sides of (15) at  $z = 1$  leads to

$$\frac{\partial P_2}{\partial x}(1,1)\frac{\sigma_2}{1-\sigma_2} + \frac{\partial P_2}{\partial z}(1,1) = \frac{(1-\rho_T) \left[ \begin{array}{l} (1-\sigma_2)B_2'(1)Y''(1) \\ -Y'(1)(1-Y'(1))\{B_2''(1) - 2B_2'(1)(B_2'(1) - \sigma_2)\} \end{array} \right]}{2(1-Y'(1))^2 B_2'(1)^2}.$$

Here  $[\partial P_2/\partial x](1,1)$  and  $[\partial P_2/\partial z](1,1)$  equal  $A_2'(1)$  and  $E[u_2]$  respectively. In the right-hand side, we still need the first and second derivative of  $Y(z)$  at  $z = 1$ .  $Y'(1)$  was already calculated in (17) and in a similar manner  $Y''(1)$  can be derived. By substituting these expressions and by introducing  $\rho_1$ ,  $\rho_2$  and  $\rho_T$ , we finally find for the mean class-2 system content:

$$\begin{aligned} E[u_2] = & \frac{\rho_2}{(1-\rho_1)(1-\rho_T)} - \frac{(1-2\sigma_1)\rho_1\rho_2}{(1-\sigma_1)(1-\rho_1)(1-\rho_T)} - \frac{(1-2\sigma_2)\rho_2\rho_T}{(1-\sigma_2)(1-\rho_T)} \\ & + \frac{\rho_2 B_1''(1)}{2(1-\sigma_1)^2(1-\rho_1)(1-\rho_T)} + \frac{\frac{\partial^2 B}{\partial z_1 \partial z_2}(1,1)}{(1-\sigma_1)(1-\sigma_2)(1-\rho_T)} + \frac{B_2''(1)}{2(1-\sigma_2)^2(1-\rho_T)}. \end{aligned} \quad (19)$$

Finally, the mean total system content  $E[u_T]$  can be calculated in a similar manner from expression (16). Obviously,  $E[u_T]=E[u_1]+E[u_2]$  must hold, which directly leads to the following expression for the mean total system content:

$$\begin{aligned} E[u_T] = & \frac{\rho_T}{1-\rho_T} - \frac{(1-2\sigma_1)\rho_1\rho_T}{(1-\sigma_1)(1-\rho_T)} - \frac{(1-2\sigma_2)\rho_2\rho_T}{(1-\sigma_2)(1-\rho_T)} \\ & + \frac{B_1''(1)}{2(1-\sigma_1)^2(1-\rho_T)} + \frac{\frac{\partial^2 B}{\partial z_1 \partial z_2}(1,1)}{(1-\sigma_1)(1-\sigma_2)(1-\rho_T)} + \frac{B_2''(1)}{2(1-\sigma_2)^2(1-\rho_T)}. \end{aligned} \quad (20)$$

In general, an explicit expression can be obtained for any moment of the class-1 and the total system content. The moments of the *class-1* system content are calculated either from expression (14) - as was the case for the mean system content - or from the functional equation (5) with  $x_2 = z_2 = 1$  or from a combination of both. The reason that we need the functional equation to calculate higher moments is because of the appearance of mixed partial derivatives with respect to  $x_1$  and  $z_1$  of the  $P$ -function when taking higher-order derivatives of (14). These mixed partial derivatives have no direct relation to the (marginal) moments of either the class-1 packet arrival process or the class-1 system content, and are therefore to be eliminated from the results. This, in turn, can be done by expressing the mixed partial derivatives in terms of other known derivatives by direct calculation from the functional equation (5). The moments of the *total* system content are calculated either from expression (16) or

from the functional equation (5) with  $z_1 = z_2 = z$  or from a combination of both. We finally note that using the functional equation to calculate the required mixed partial derivatives is possible in both cases because either the function  $P(0, 0, 1 - \sigma_2 + \sigma_2 x_2 z_2, z_2)$  becomes the constant  $P_1(0, 0)$  in (5) when substituting  $x_2 = z_2 = 1$  (for the class-1 system content) or the term in  $P(0, 0, 1 - \sigma_2 + \sigma_2 x_2 z_2, z_2)$  disappears when substituting  $z_1 = z_2 = z$  (for the total system content). Unfortunately neither is the case when one wants to calculate the higher moments of the *class-2* system content. The calculation of these moments is still an open issue at the moment.

## 4 Packet delay

We study the delay experienced by class-1 and class-2 packets in the system in this section. The delay of a packet is defined as the number of slots between the end of the packet's slot of arrival and the end of its departure slot (thus excluding its arrival slot and including its departure slot). Within each class, we assume that packets are transmitted in the order of their arrival. Class-1 packets obviously still have HOL priority over class-2 packets. We analyze the class-1 and class-2 packet delays separately in the remainder.

### 4.1 Class-1 packet delay

The analysis of the class-1 packet delay is rather easy once the observation is made that class-1 packets are not influenced by class-2 packets in the system, due to the HOL priority scheduling discipline. The expression for  $D_1(z)$ , the pgf of the class-1 packet delay in the steady state, is directly related to the pgf  $P_1(1, z)$  of the system content of class 1 at the beginning of a random slot, as follows (see [11]):

$$D_1(z) = \frac{P_1(1, z) - 1 + \rho_1}{\rho_1}. \quad (21)$$

This thus allows us to derive the moments of the class-1 packet delay, which can be expressed in terms of the derivatives of  $D_1(z)$  at  $z = 1$ , as functions of the moments of the class-1 system content. We have shown in the previous section how the latter are calculated by taking the required partial derivatives of expressions (5) and (14) for the arguments equal to 1. E.g. the mean class-1 packet delay  $E[d_1]$  is calculated by taking the first derivative of both sides of (21), yielding

$$E[d_1] = D_1'(1) = \frac{1}{\rho_1} \frac{\partial P_1}{\partial z}(1, 1) = \frac{E[u_1]}{\rho_1} \quad (22)$$

This is also a direct consequence of Little's law. Substituting (18) in this expression yields

$$E[d_1] = \frac{1}{1 - \rho_1} - \frac{\rho_1(1 - 2\sigma_1)}{(1 - \sigma_1)(1 - \rho_1)} + \frac{B_1''(1)}{2(1 - \sigma_1)^2\rho_1(1 - \rho_1)}. \quad (23)$$

Higher moments of the class-1 packet delay can be calculated as well.

## 4.2 Class-2 packet delay

The analysis of the steady-state class-2 packet delay is more involved, because the delay of a class-2 packet is also influenced by later arriving class-1 packets, due to the HOL priority discipline. We tag a random class-2 packet and denote it by  $Q_2$ . We denote the slot during which  $Q_2$  arrives by  $S_2$ . We first make the following key observation, which highly simplifies the analysis: if a class-1 packet is transmitted before  $Q_2$ , all packets of the same message of this class-1 packet are transmitted before  $Q_2$  as well. Indeed, only other class-1 packets can be transmitted between the transmissions of two randomly chosen packets of a *same* class-1 message.

Furthermore, we denote the numbers of class-1 and class-2 arrivals during slot  $S_2$  by  $a_1^*$  and  $a_2^*$  respectively and the total system content at the beginning of the following slot by  $u_T^*$ . Furthermore, let  $r_2$  indicate the number of packets arriving during slot  $S_2$  and to be transmitted after packet  $Q_2$ . Before writing down an expression for and analyzing the delay of  $Q_2$ , we first concentrate on the *virtual delay*  $w_2$  of  $Q_2$ . This virtual delay is defined as the delay when no *new messages* arrive after slot  $S_2$ .  $w_2$  then equals

$$w_2 = u_T^* - r_2 + \sum_{i=1}^{a_1^*} l_{1,i}^+, \quad (24)$$

with  $l_{1,i}^+$  the number of packets arriving after slot  $S_2$  of the message which generated the  $i$ -th class-1 arrival during slot  $S_2$ . The virtual delay thus equals the system content just after slot  $S_2$  and to be transmitted no later than  $Q_2$  and packets arriving after slot  $S_2$  of class-1 messages which were already generating a packet during slot  $S_2$ . Note that due to the geometric distribution of the class-1 message lengths, the  $l_{1,i}^+$  are all independent of the system state just after slot  $S_2$  and are all geometrically distributed with parameter  $\sigma_1$ . They thus have the following (common) pgf:

$$L_1^+(z) = \frac{1 - \sigma_1}{1 - \sigma_1 z}. \quad (25)$$

Denoting the joint pgf of  $a_1^*$ ,  $r_2$  and  $u_T^*$  by  $Q(x_1, x_2, z)$ , i.e.,

$$Q(x_1, x_2, z) \triangleq \mathbf{E} \left[ x_1^{a_1^*} x_2^{r_2} z^{u_T^*} \right],$$

expression (24) leads to the pgf of  $w_2$ :

$$W_2(z) \triangleq \mathbf{E}[z^{w_2}] = Q(L_1^+(z), 1/z, z). \quad (26)$$

We will calculate  $Q(x_1, x_2, z)$  at the end of this section. We first look for a relation between the delay  $d_2$  and the virtual delay  $w_2$  of  $Q_2$ . Obviously, the virtual delay is part of the delay. During the transmission of a certain packet, say  $P$ , belonging to the virtual delay workload, new class-1 messages may be generated. The packets of these messages all add to the delay of  $Q_2$ . During the transmission of these class-1 messages also new class-1 messages may arrive, which further add to the delay of  $Q_2$ , etc. The total number of all packets of all these messages (including packet  $P$  itself) is called the *sub-busy period* initiated by  $P$ . Summarizing, we can write

$$d_2 = \sum_{i=1}^{w_2-1} v_{1,i} + 1, \quad (27)$$

with  $v_{1,i}$  the sub-busy period added by the  $i$ -th packet of the virtual delay workload. Note that these  $v_{1,i}$ 's are all i.i.d. with pgf  $V_1(z)$ . By  $z$ -transforming expression (27), we then obtain

$$D_2(z) \triangleq \mathbf{E}[z^{d_2}] = \frac{zW_2(V_1(z))}{V_1(z)}.$$

Using (26), we find

$$D_2(z) = \frac{zQ(L_1^+(V_1(z)), 1/V_1(z), V_1(z))}{V_1(z)}. \quad (28)$$

Note that the  $v_{1,i}$ 's can be expressed as

$$v_{1,i} = 1 + \sum_{m=1}^{b_{1,i}} \sum_{n=1}^{l_{1,i}^{(m)}} v_{1,i}^{(m,n)}, \quad (29)$$

with  $b_{1,i}$  the number of new class-1 messages generated during the transmission of the  $i$ -th packet of the virtual delay workload,  $l_{1,i}^{(m)}$  the message length of the  $m$ -th message of  $b_{1,i}$  and  $v_{1,i}^{(m,n)}$  the sub-busy

period initiated by the  $n$ -th packet of the  $m$ -th message of  $b_{1,i}$ . Indeed, a sub-busy period initiated by a packet consists of the transmission slot of that packet and the sub-busy periods of all packets of all messages that are generated during that slot. Note that the  $v_{1,i}^{(m,n)}$  are i.i.d. having the same pgf as the  $v_{1,i}$ 's, i.e.,  $V_1(z)$ . Expression (29) then leads to the following functional equation for  $V_1(z)$ :

$$V_1(z) = zB_1(L_1(V_1(z))). \quad (30)$$

Going back to expression (28), we are left with the task to calculate  $Q(x_1, x_2, z)$ . This is written as function of the corresponding joint probability mass function (pmf) as follows:

$$Q(x_1, x_2, z) = \sum_{i=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=1}^{\infty} \text{Prob}[a_1^* = i, r_2 = l, u_T^* = m] x_1^i x_2^l z^m. \quad (31)$$

The pmf can be determined by conditioning on the value of the random variable  $a_2^*$ :

$$\begin{aligned} \text{Prob}[a_1^* = i, r_2 = l, u_T^* = m] &= \sum_{n=l+1}^{m-i} \text{Prob}[r_2 = l | a_1^* = i, a_2^* = n, u_T^* = m] \\ &\quad \times \text{Prob}[a_1^* = i, a_2^* = n, u_T^* = m], \end{aligned} \quad (32)$$

for  $m \geq 1$  and  $i + l \leq m - 1$ . The quantities on the right-hand side of this expression can be found as follows. First, we note that

$$\text{Prob}[r_2 = l | a_1^* = i, a_2^* = n, u_T^* = m] = \frac{1}{n}, \quad \text{for } 0 \leq l \leq n - 1, \quad (33)$$

due to the fact that  $Q_2$  is an arbitrary class-2 packet in a batch of  $n$  class-2 packets. Next, in order to derive  $\text{Prob}[a_1^* = i, a_2^* = n, u_T^* = m]$  we first note that this quantity does not correspond to *the fraction of slots* which have  $i$  class-1 packet arrivals,  $n$  class-2 packet arrivals and a total system content of  $m$  just after that slot, but rather to *the fraction of class-2 packets* that arrive in such a slot. This is in fact a manifestation of the well-known renewal theory paradox. Since each such slot contains exactly  $n$  class-2 packet arrivals and the arbitrary packet  $Q_2$  could be any of those, it is clear that

$$\text{Prob}[a_1^* = i, a_2^* = n, u_T^* = m] = \frac{np_T(i, n, m)}{\rho_2}, \quad \text{for } i \geq 0, n \geq 1, m \geq i + n, \quad (34)$$

with  $p_T(i, n, m)$  the steady-state probability of having  $i$  class-1 and  $n$  class-2 packet arrivals in an

arbitrary slot and a total system content of  $m$  just after this slot. Obviously, the joint pgf  $P_T(x_1, x_2, z)$  defined in (13) is related to this joint pmf as follows:

$$P_T(x_1, x_2, z) = \sum_{i=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} p_T(i, n, m) x_1^i x_2^n z^m. \quad (35)$$

Equations (32)-(34) can now be combined to

$$\text{Prob}[a_1^* = i, r_2 = l, u_T^* = m] = \frac{1}{\rho_2} \sum_{n=l+1}^{m-i} p_T(i, n, m), \quad \text{for } m \geq 1 \text{ and } i + l \leq m - 1.$$

Using this in (31) yields

$$Q(x_1, x_2, z) = \frac{1}{\rho_2} \sum_{m=1}^{\infty} \sum_{i=0}^{m-1} \sum_{l=0}^{m-i-1} \sum_{n=l+1}^{m-i} p_T(i, n, m) x_1^i x_2^l z^m.$$

Reordering the summations, working out the sum over  $l$  and using (35), we get

$$Q(x_1, x_2, z) = \frac{P_T(x_1, 1, z) - P_T(x_1, x_2, z)}{\rho_2(1 - x_2)}.$$

Finally, using this expression in expression (28) provides us with an expression for  $D_2(z)$ :

$$D_2(z) = \frac{z [P_T(L_1^+(V_1(z)), 1, V_1(z)) - P_T(L_1^+(V_1(z)), 1/V_1(z), V_1(z))]}{\rho_2(V_1(z) - 1)}, \quad (36)$$

with  $L_1^+(z)$  as given in (25) and  $V_1(z)$  implicitly defined in (30). This expression gives us the opportunity to calculate the moments of the class-2 packet delay as functions of (partial) derivatives of the  $P$ -function for  $z_1 = z_2 = z$  and derivatives of  $V_1(z)$ , all evaluated in  $x_1 = x_2 = z = 1$ . We have argued in the previous section that the derivatives of the  $P$ -function for  $z_1 = z_2 = z$  can be calculated. Calculating derivatives of  $V_1(z)$  in  $z = 1$  is possible as well due to the knowledge that  $V_1(1) = 1$ , since  $V_1(z)$  is a pgf. E.g. calculating the first derivative of both sides of (30) and solving for  $V_1'(1)$  yields

$$V_1'(1) = \frac{1}{1 - \rho_1}.$$

We demonstrate the calculation of the mean class-2 packet delay. We find

$$E[d_2] = D_2'(1)$$

$$= \frac{1}{\rho_2(1-\rho_1)} \left[ \frac{\sigma_1}{1-\sigma_1} \frac{\partial^2 P_T}{\partial x_1 \partial x_2}(1,1,1) - \rho_1 \frac{\partial P_T}{\partial x_2}(1,1,1) - \frac{1}{2} \frac{\partial^2 P_T}{\partial x_2^2}(1,1,1) + \frac{\partial^2 P_T}{\partial x_2 \partial z}(1,1,1) \right]. \quad (37)$$

The first three partial derivatives in this expression are related to the moments of the arrival process and are respectively given by

$$\begin{aligned} \frac{\partial^2 P_T}{\partial x_1 \partial x_2}(1,1,1) &= \frac{\partial^2 A}{\partial x_1 \partial x_2}(1,1), \\ \frac{\partial P_T}{\partial x_2}(1,1,1) &= A'_2(1), \\ \frac{\partial^2 P_T}{\partial x_2^2}(1,1,1) &= A''_2(1). \end{aligned}$$

These can thus be calculated from expression (6). The fourth partial derivative can be calculated from the functional equation (5) with  $z_1 = z_2 = z$ , which yields

$$\begin{aligned} \frac{\partial^2 P_T}{\partial x_2 \partial z}(1,1,1) &= \rho_2(1 + E[u_T]) + \frac{\frac{\partial^2 B}{\partial z_1 \partial z_2}(1,1)}{(1-\sigma_2\sigma_1)(1-\sigma_2)} + \frac{B''_2(1)}{(1-\sigma_2)^2(1+\sigma_2)} \\ &\quad - \frac{\rho_2[\rho_T\sigma_2^3 - ((1-2\sigma_1)\rho_T + \rho_1)\sigma_2^2 - (\rho_2(1+\sigma_1) + \rho_T)\sigma_2 + \rho_T - \rho_1\sigma_1]}{(1-\sigma_2\sigma_1)(1-\sigma_2)(1+\sigma_2)}. \end{aligned}$$

Substituting the expressions of these (mixed) partial derivatives and expression (20) of  $E[u_T]$  in expression (37), we obtain the following expression for the mean class-2 packet delay

$$\begin{aligned} E[d_2] &= \frac{1}{(1-\rho_1)(1-\rho_T)} - \frac{(1-2\sigma_1)\rho_1}{(1-\sigma_1)(1-\rho_1)(1-\rho_T)} - \frac{(1-2\sigma_2)\rho_T}{(1-\sigma_2)(1-\rho_T)} \\ &\quad + \frac{B''_1(1)}{2(1-\sigma_1)^2(1-\rho_1)(1-\rho_T)} + \frac{\frac{\partial^2 B}{\partial z_1 \partial z_2}(1,1)}{(1-\sigma_1)(1-\sigma_2)\rho_2(1-\rho_T)} + \frac{B''_2(1)}{2(1-\sigma_2)^2\rho_2(1-\rho_T)}. \end{aligned} \quad (38)$$

From expressions (18), (19), (23) and (38) it can be seen that the mean system content and packet delay of both classes fulfill Little's theorem, i.e.,  $E[u_j] = \rho_j E[d_j]$  for  $j = 1, 2$ . Note however, that the developments in this section are necessary if one wants to obtain higher order moments of the class-1 and class-2 packet delay. The calculations of these higher moments include the calculation of a numerous number of partial (mixed) derivatives (as already explained in section 3). As a result, the obtained expressions of these higher order moments are quite large and are thus not shown here. In order to show that it is possible to calculate these higher order moments though, we will show some plots of the variances of the packet delays in the next section.



## 5 Numerical examples

### 5.1 Poisson distributed numbers of per-slot newly generated messages

We illustrate our findings by means of some numerical examples. We firstly consider the case that class-1 and class-2 messages are both generated according to independent Poisson processes with means  $\rho_1(1 - \sigma_1)$  and  $\rho_2(1 - \sigma_2)$  respectively. We thus have

$$B(z_1, z_2) = e^{\rho_1(1-\sigma_1)(z_1-1)} e^{\rho_2(1-\sigma_2)(z_2-1)}.$$

The pgf of the compound packet arrival process of both classes can be derived from (7), resulting in

$$A(z_1, z_2) = e^{\rho_1(z_1-1)} e^{\rho_2(z_2-1)}. \quad (39)$$

The numbers of class- $j$  packet arrivals per slot in the steady state thus also have a Poisson distribution, with parameter  $\rho_j$ , for  $j = 1, 2$ . Furthermore, we define  $\alpha$  as the fraction of the class-1 load in the total load, i.e., as  $\rho_1/\rho_T$ .

In this subsection, we will firstly concentrate on the mean packet delays of both classes and the mean packet delay of an arbitrary packet. Obviously, the means of the system contents can be illustrated in a similar manner. Plots of the variances of the packet delays will be shown at the end of the subsection. Note that the mean packet delay of an arbitrary packet can be calculated either from the expressions of the mean delays of both classes (namely it equals  $\alpha E[d_1] + (1 - \alpha)E[d_2]$ ), or from the total mean system content using Little's law. Note further that since the total mean system content is independent of the scheduling discipline - as long as it is work-conserving - the mean packet delay of an arbitrary packet is as well. The curves for the mean packet delay of an arbitrary packet in the figures in this section are thus valid for any work-conserving scheduling discipline (and thus e.g. also for FIFO). A similar reasoning leads to the observation that the mean delay of an arbitrary packet is independent of  $\alpha$  as well.

Figure 1 shows the mean packet delays of both classes and the mean packet delay of an arbitrary packet as functions of the total load for  $\alpha = 0.25, 0.5$  and  $0.75$  and for  $\sigma_1 = \sigma_2 = 0.9$ . We observe that the influence of a priority scheduling discipline is quite large. The mean delay of class-1 packets reduces considerably compared to the mean delay of an arbitrary packet. The price to pay is obviously a larger mean class-2 packet delay. Also note that it follows from this figure that increasing the fraction

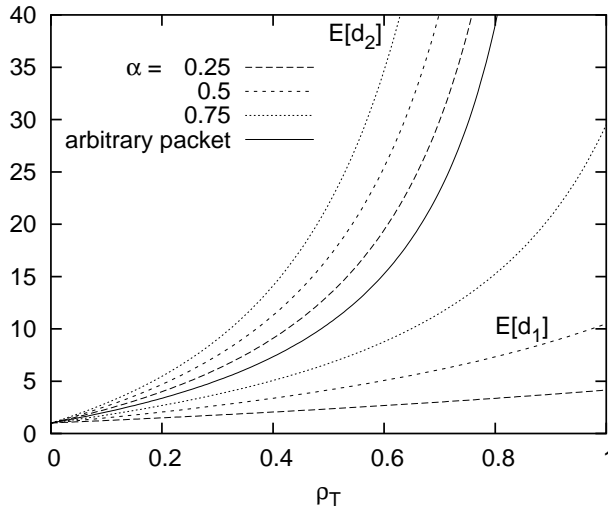


Figure 1: Mean delays versus the total load for various values of  $\alpha$

of the class-1 load in the overall traffic mix increases the mean delay of class-1 *and* class-2 packets.

In Figure 2, we depict the mean packet delays of both classes and the mean packet delay of an arbitrary packet as functions of the total load for  $\alpha = 0.25$  and for  $\sigma_1 = \sigma_2 = \sigma = 0, 0.5, 0.9$  and  $0.99$ . These latter values correspond to mean message lengths of 1, 2, 10 and 100 respectively. In order to not overload the figure, we split it in two parts: Figure 2a. shows the mean class-1 packet delay while Figure 2b. shows the mean class-2 packet delay and the mean delay of an arbitrary packet. One can easily see the influence of correlation in the arrival process: for  $\sigma = 0$ , we have the uncorrelated case, i.e., all messages consist of exactly one packet and the corresponding packet arrival process is uncorrelated from slot to slot. When  $\sigma$  increases, i.e., when the arrival process becomes more correlated, the mean delays increase as well. The increase of the mean class-2 packet delay is however much larger than the increase of the mean class-1 packet delay, as can be seen by comparing both parts of the figure. Correlation in the arrival process thus has a larger impact on the delay of low-priority packets.

Figure 3 depicts the mean class-1 and class-2 packet delays as functions of the total load for  $\alpha = 0.25$ ,  $\sigma_1 = 0.9$  and  $\sigma_2 = 0, 0.5, 0.9$  and  $0.99$ . Since it is clear from the previous two plots that the curves for the mean delay of an arbitrary packet are always near to the corresponding curves of the mean class-2 packet delay, we do not show the former in this figure (or in the next figures). As expected the mean class-1 (high-priority) packet delay is independent of the correlation in the class-2 (low-priority) packet arrival process. The mean class-2 packet delay is obviously largely influenced by this correlation in the class-2 packet arrival process.

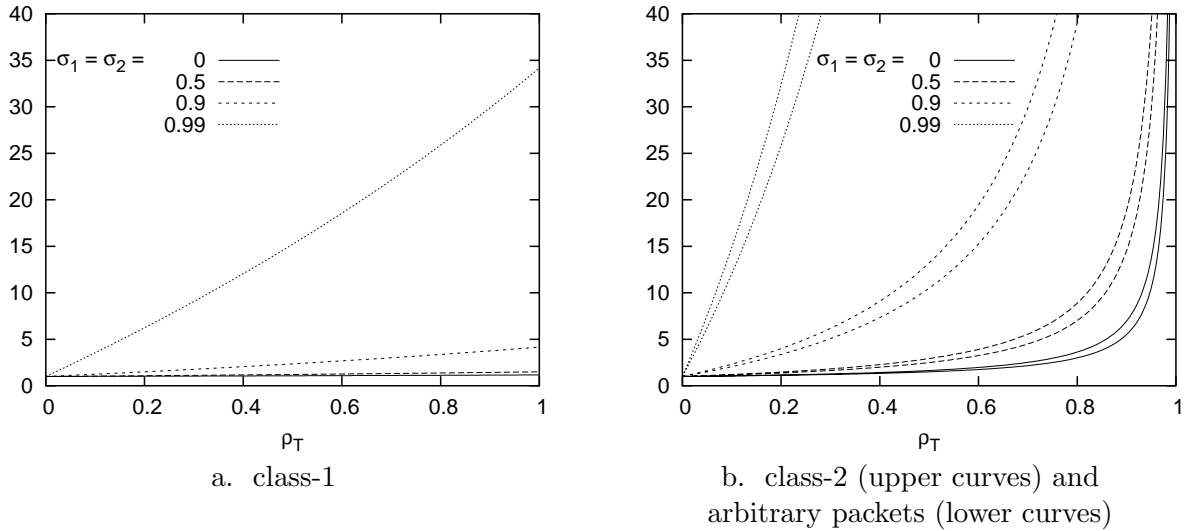


Figure 2: Mean delay versus the total load for various values of  $\sigma_1 = \sigma_2$

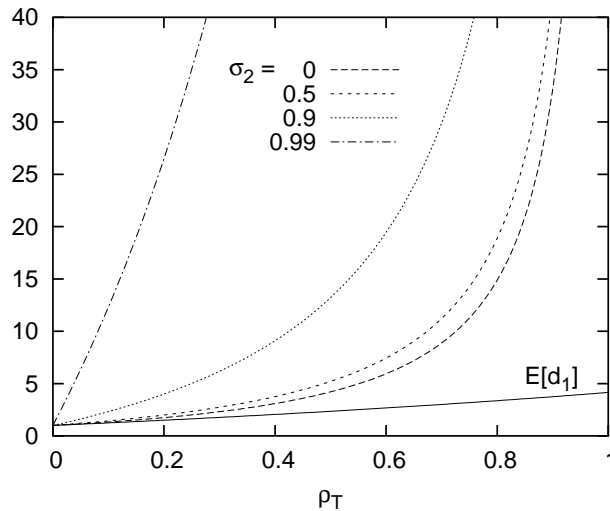


Figure 3: Mean delays versus the total load for various values of  $\sigma_2$

Figures 4a. and 4b. show the mean class-1 and class-2 packet delays respectively as functions of the total load for  $\alpha = 0.25$ ,  $\sigma_2 = 0.9$  and  $\sigma_1 = 0, 0.5, 0.9$  and  $0.99$ . First note that Figure 4a. is identical to Figure 2a., since in both figures the arrival processes of the class-1 packets are identical. The most important figure is however Figure 4b. This figure shows the influence of the correlation in the *high-priority* packet arrival process on the mean *low-priority* delay. Thus the large increase of the mean class-2 packet delay in Figure 2b. is not only the result of the correlation in the low-priority packet arrival process; it is clear from Figure 4b. that the mean packet delay of the low-priority traffic is also highly influenced by (the correlation in) the high-priority arrival process.

Finally, we show some plots of the variances of the class-1 and class-2 packet delays. This is not

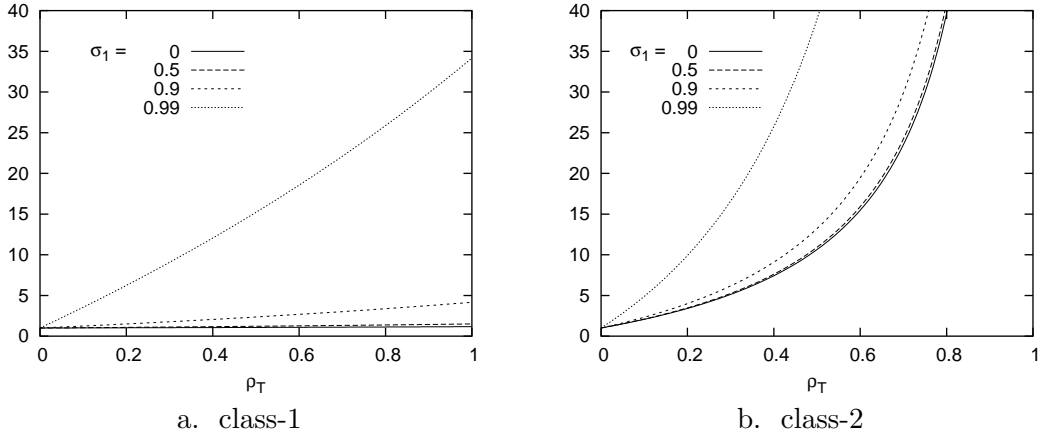


Figure 4: Mean delay versus the total load for various values of  $\sigma_1$

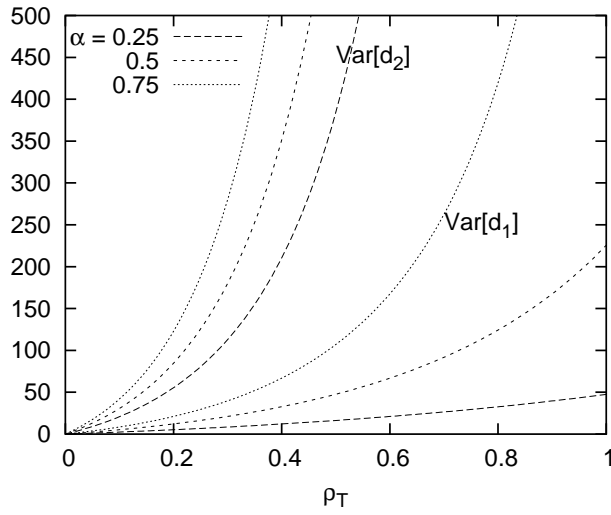


Figure 5: Variance of delays versus the total load for various values of  $\alpha$

only to show the influence of the parameters on these performance measures, but also to show that it is possible to effectively calculate these variances with the techniques demonstrated in this paper. Figure 5 shows the variances of the class-1 and class-2 packet delays as functions of the total load for  $\alpha = 0.25, 0.5$  and  $0.75$  and for  $\sigma_1 = \sigma_2 = 0.9$ . Finally, Figure 6 shows the variance of the low-priority packet delay as function of the total load for  $\alpha = 0.25, \sigma_2 = 0.9$  and  $\sigma_1 = 0, 0.5, 0.9$  and  $0.99$ . From this last figure, it is again obvious that the low-priority performance highly depends on the characteristics of and the correlation in the high-priority arrival stream.

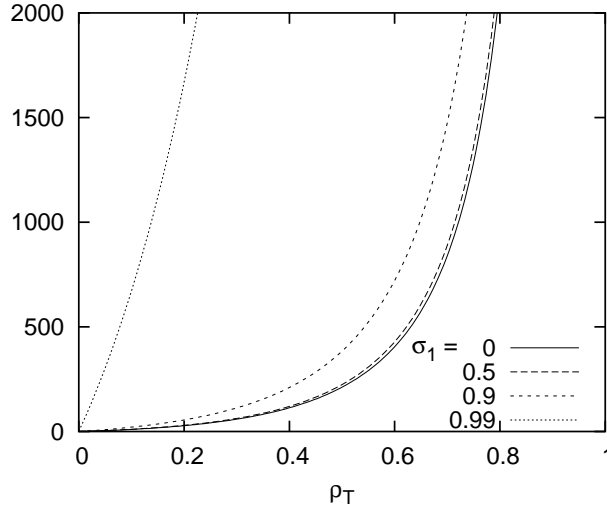


Figure 6: Variance of class-2 delay versus the total load for various values of  $\sigma_1$

## 5.2 Correlated numbers of newly generated class-1 and class-2 messages

In a last example, we show the influence of correlation between the numbers of newly generated class-1 and class-2 messages in a slot. We therefore assume the following message generation process in this subsection: the number of class- $j$  generations during a slot is Bernoulli distributed with parameter  $\lambda_j$ .  $B'_j(1)$  thus equals  $\lambda_j$  and we have that  $\lambda_j = \rho_j(1 - \sigma_j)$ . The joint pgf  $B(z_1, z_2)$  of the numbers of newly generated messages of both classes is given by

$$B(z_1, z_2) = 1 - \lambda_1 - \lambda_2 + q_{12} + (\lambda_1 - q_{12})z_1 + (\lambda_2 - q_{12})z_2 + q_{12}z_1z_2,$$

with  $q_{12}$  a parameter. The correlation factor  $\rho_{b_1b_2}$  between the numbers of generated class-1 and class-2 messages in a slot is given by

$$\rho_{b_1b_2} = \frac{q_{12} - \lambda_1\lambda_2}{\sqrt{\lambda_1\lambda_2(1 - \lambda_1)(1 - \lambda_2)}}.$$

By varying  $q_{12}$ , this correlation factor can be varied, while keeping the mean numbers of generated messages of both classes constant ( $\lambda_1$  and  $\lambda_2$  respectively). Positive correlation may occur in practice when a new user for example starts generating high- and low-priority packets at the same time. Since the delay of class-1 packets is independent of  $\rho_{b_1b_2}$ , as indicated by equations (14) and (21), we focus on the class-2 packet delay in this example. In Figure 7, we show the mean delay of class-2 packets versus the total load, for  $\sigma_1 = \sigma_2 = 0.9$  and  $\alpha = 0.25$  and  $0.75$  respectively. For both cases, we

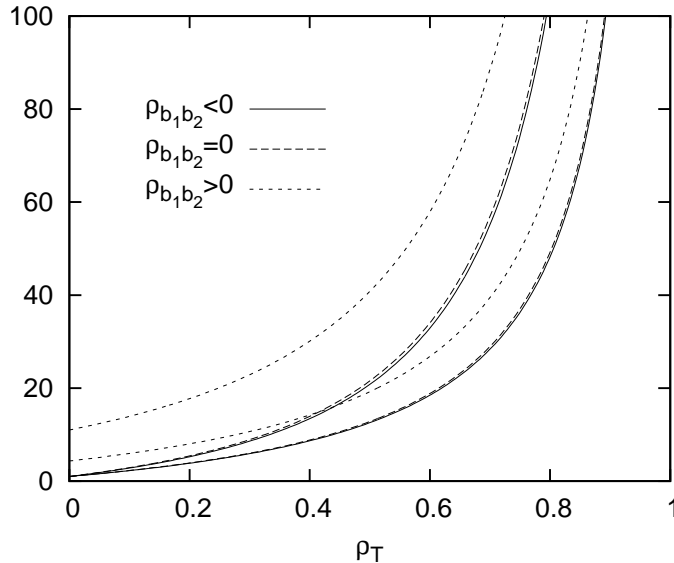


Figure 7: Mean packet delay of class 2 versus the total load for negative, zero and positive correlation between the number of newly generated class-1 and class-2 messages. For each correlation factor two curves are drawn:  $\alpha = 0.25$  for the lower curves and  $\alpha = 0.75$  for the upper

consider 3 values for the correlation factor: its minimal value (when we choose  $q_{12}=0$ ), the value zero (for  $q_{12} = \lambda_1 \lambda_2$ ) and its maximal value (when  $q_{12}=\min(\lambda_1, \lambda_2)$ ). We see that the influence of the correlation factor is limited for negative correlation, while it is significant for positive correlation. The reason for the higher mean class-2 delay for positive correlation between the numbers of new messages of both classes is the increasing probability that messages of class 1 are arriving at the same time as messages of class 2. We further see from Figure 7 that a higher value of  $\alpha$  leads to a higher mean class-2 delay, as already observed before in Figure 1.

## 6 Conclusions

In this paper, we analyzed a discrete-time queue with a HOL priority scheduling discipline and two priority classes. Variable-sized messages of both priority classes arrive to the system as trains of fixed-sized packets, i.e., the packets of a message arrive at the rate of one packet per slot. The messages are assumed to have geometrically distributed lengths. Obviously, the arrival process is correlated, which highly complicates the analysis. Using probability generating functions, we have shown that explicit closed-form expressions for the mean values of the system contents and packet delays of both classes can be derived, as well as higher moments for the packet delays of both classes. We have shown the influence of all parameters of the arrival process on the mean and variance of the packet delay. The

most important conclusion is that the (high- and) low-priority packet delay is highly influenced by the correlation in the high-priority packet arrival process. This correlation thus has to be taken into account if one wants to estimate the mean packet delay (or the buffer size) of both classes.

## Acknowledgment

The first author is a Postdoctoral Fellow with the Fund for Scientific Research, Flanders (F.W.O.-Vlaanderen), Belgium.

## References

- [1] H. Bruneel. Packet delay and queue length for statistical multiplexers with low-speed access lines. *Computer Networks and ISDN Systems*, 25(12):1267–1277, 1993.
- [2] H. Bruneel. Calculation of message delays and message waiting times in switching elements with slow access lines. *IEEE Transactions on Communications*, 42(2/3/4):255–259, 1994.
- [3] B. Choi, D. Choi, Y. Lee, and D. Sung. Priority queueing system with fixed-length packet-train arrivals. *IEE Proceedings-Communications*, 145(5):331–336, 1998.
- [4] S. De Vuyst, S. Wittevrongel, and H. Bruneel. Statistical multiplexing of correlated variable-length packet trains: an analytic performance study. *Journal of the Operational Research Society*, 52(3):318–327, 2001.
- [5] N. Jaiswal. *Priority queues*. Academic Press, New York, 1968.
- [6] K. Laevens and H. Bruneel. Discrete-time multiserver queues with priorities. *Performance Evaluation*, 33(4):249–275, 1998.
- [7] M. Mehmet Ali and X. Song. A performance analysis of a discrete-time priority queueing system with correlated arrivals. *Performance Evaluation*, 57(3):307–339, 2004.
- [8] M. Neuts. Probabilistic modelling requires a certain imagination. In *Proceedings of the 3rd International Conference on Teaching Statistics (ICOTS 3)*, pages 122–131, Dunedin, 1990.
- [9] H. Takagi. *Queueing analysis: a foundation of performance evaluation, volume 1: vacation and priority systems, part 1*. North-Holland, 1991.

- [10] T. Takine, B. Sengupta, and T. Hasegawa. An analysis of a discrete-time queue for broadband ISDN with priorities among traffic classes. *IEEE Transactions on Communications*, 42(2-4):1837–1845, 1994.
- [11] B. Vinck and H. Bruneel. A note on the system contents and cell delay in FIFO ATM-buffers. *Electronics Letters*, 31(12):952–954, 1995.
- [12] J. Walraevens, B. Steyaert, and H. Bruneel. Delay characteristics in discrete-time GI-G-1 queues with non-preemptive priority queueing discipline. *Performance Evaluation*, 50(1):53–75, 2002.
- [13] J. Walraevens, B. Steyaert, and H. Bruneel. Performance analysis of a single-server ATM queue with a priority scheduling. *Computers & Operations Research*, 30(12):1807–1829, 2003.
- [14] S. Wittevrongel. Discrete-time buffers with variable-length train arrivals. *Electronics Letters*, 34(18):1719–1721, 1998.
- [15] S. Wittevrongel and H. Bruneel. Correlation effects in ATM queues due to data format conversions. *Performance Evaluation*, 32(1):35–56, 1998.