

# A Discrete-Time HOL Priority Queue with Multiple Traffic Classes

Joris Walraevens, Bart Steyaert, Marc Moeneclaey and Herwig Bruneel

Ghent University, Department TELIN (TW07)  
Sint-Pietersnieuwstraat 41, B-9000 Gent, Belgium.  
jw@telin.UGent.be

**Abstract.** Priority scheduling for packets is a hot topic, as interactive (voice, video) services are being integrated in existing data networks. In this paper, we consider a discrete-time queueing system with non-preemptive (or Head-Of-the-Line) priority scheduling and a general number of priority classes. Packets of variable length arrive in the queueing system. We derive expressions for the probability generating functions of the packet delays. From these functions, some performance measures (such as moments and approximate tail probabilities) are calculated. We apply the theoretical results to a queue that handles arriving multimedia traffic.

## 1 Introduction

In recent years, there has been much interest in incorporating multimedia applications in packet-based networks (e.g. IP networks). Different types of traffic need different QoS standards. For real-time interactive applications, it is important that mean delay and delay-jitter are bound, while for non real-time applications, the loss ratio is the restrictive quantity. In order to guarantee acceptable delay boundaries to delay-sensitive traffic (such as voice/video), several scheduling schemes – for switches, routers, ... – have been proposed and analyzed, each with their own specific algorithmic and computational complexity. The (strict) priority scheduling is the most drastic one. With this scheduling, as long as delay-sensitive (or high-priority) packets are present in the queueing system, they will be served first. Delay-insensitive packets can thus only be transmitted when no delay-sensitive traffic is present in the system. Clearly, this is the most rigorous way to meet the QoS constraints of delay-sensitive traffic (and thus the scheduling with the most disadvantageous consequences to the delay-insensitive traffic), but also the easiest to implement. In the related literature, there have been a number of contributions with respect to HOL priority scheduling (see e.g. [1–7]).

In this paper, we focus on the effect of a non-preemptive or HOL (Head-Of-the-Line) priority scheduling discipline on the performance of a queue with multiple traffic classes. More delay-sensitive traffic is assumed to have non-preemptive priority over traffic with more flexible delay constraints, i.e., when the server

becomes idle, a packet of the most delay-sensitive traffic that is available is scheduled for service next. Newly arriving packets cannot interrupt the transmission of a packet that has already commenced, whatever their priority level. The transmission times of the packets are assumed to be generally distributed and class-dependent (which reflects the case where different classes represent different applications). We will demonstrate that an analysis based on probability generating functions (pgf's) is extremely suitable for modeling this type of buffers with a priority scheduling discipline. From these pgf's, we calculate closed-form expressions for some interesting performance measures.

## 2 Mathematical Model

We consider a discrete-time single-server queueing system with infinite buffer space. Time is assumed to be slotted. There are  $M$  types of traffic arriving in the system. We denote the number of packet arrivals of class- $j$  during slot  $k$  by  $a_{j,k}$  ( $j = 1, \dots, M$ ). All types of packet arrivals are assumed to be i.i.d. from slot-to-slot and the number of per-slot arrivals are characterized by the joint pgf

$$A(\mathbf{z}) \triangleq \mathbb{E} \left[ \prod_{j=1}^M z_j^{a_{j,k}} \right],$$

with  $\mathbf{z}$  defined as a vector with elements  $z_j$  ( $j = 1, \dots, M$ ). We define the marginal pgf's of the arrivals from class- $j$  during a slot by

$$A_j(z) \triangleq \mathbb{E}[z^{a_{j,k}}] = A(\mathbf{z}) \Big|_{z_j=z, z_i=1, i \neq j}.$$

We furthermore denote the arrival rate of class- $j$  ( $j = 1, \dots, M$ ) by  $\lambda_j = A'_j(1)$  and the total arrival rate by  $\lambda_T = \sum_{j=1}^M \lambda_j$ . The service times of the class- $j$  packets are assumed to be i.i.d. and are characterized by the probability mass function  $s_j(m)$ ,  $m \geq 1$ , and pgf  $S_j(z) = \sum_{m=1}^{\infty} s_j(m)z^m$ , with  $j = 1, \dots, M$ . We furthermore denote the mean service time of a class- $j$  packet by  $\mu_j = S'_j(1)$  and define the load offered by class- $j$  packets as  $\rho_j \triangleq \lambda_j \mu_j$ . The total load is given by  $\rho_T \triangleq \sum_{j=1}^M \rho_j$ , and the equilibrium condition requires that  $\rho_T < 1$ .

The system has one server that provides for the transmission of the packets. Class- $i$  packets are assumed to have non-preemptive priority over class- $j$  packets when  $i < j$ , and within one class the service discipline is FCFS.

## 3 System Contents at Service Initiation Epochs

In order to be able to analyze the packet delay, we first analyze the system contents at the beginning of so-called start-slots, i.e., slots at the beginning of which a packet (if available) can enter the server. Note that every slot during which the system is empty is also a start-slot. We denote the system contents  $n_{j,l}$  as the number of class- $j$  ( $j = 1, \dots, M$ ) packets in the buffer at the beginning

of the  $l$ -th start-slot, including the packet being served (if any). Their joint pgf is denoted by

$$N_l(\mathbf{z}) \triangleq \mathbb{E} \left[ \prod_{j=1}^M z_j^{n_{j,l}} \right]. \quad (1)$$

The set  $\{(n_{1,l}, \dots, n_{M,l}), l \geq 1\}$  forms a Markov chain, since the arrival process is i.i.d. and the buffer solely contains entire packets at the beginning of start-slots.  $s_l^*$  denotes the service time of the packet that enters service at the beginning of start-slot  $l$  (which corresponds - by definition - to regular slot  $k$ ). We then establish the following system equations:

- If  $n_{1,l} = \dots = n_{M,l} = 0$ :

$$n_{i,l+1} = a_{i,k} \quad \text{for } i = 1, \dots, M.$$

- If  $n_{1,l} = \dots = n_{j-1,l} = 0, n_{j,l} > 0$ :

$$n_{i,l+1} = n_{i,l} - 1_{i=j} + \sum_{m=0}^{s_i^*-1} a_{i,k+m} \quad \text{for } i = 1, \dots, M,$$

for  $j = 1, \dots, M$ .  $1_X$  is the indicator function of  $X$ .

Using the system equations, we can derive a relation between  $N_l(\mathbf{z})$  and  $N_{l+1}(\mathbf{z})$ . We assume that the system is stable (implying that the equilibrium condition  $\rho_T < 1$  is satisfied) and as a result  $N_l(\mathbf{z})$  and  $N_{l+1}(\mathbf{z})$  converge both to a common steady-state value  $N(\mathbf{z})$  for  $l \rightarrow \infty$ . By taking the  $l \rightarrow \infty$  limit of the relation between  $N_l(\mathbf{z})$  and  $N_{l+1}(\mathbf{z})$ , we obtain:

$$N(\mathbf{z}) = \frac{z_1}{z_1 - S_1(A(\mathbf{z}))} \left\{ \frac{z_M A(\mathbf{z}) - S_M(A(\mathbf{z}))}{z_M} N(\mathbf{0}) + \sum_{j=2}^M \left[ \frac{S_j(A(\mathbf{z}))}{z_j} - \frac{S_{j-1}(A(\mathbf{z}))}{z_{j-1}} \right] N(\mathbf{z}_j) \right\}. \quad (2)$$

There are  $M$  quantities yet to be determined in the right hand side of equation (2), namely the functions  $N(\mathbf{z}_j)$  ( $j = 2, \dots, M$ ) and the constant  $N(\mathbf{0})$ . First, we will recursively express  $N(\mathbf{z}_m)$  ( $m = 1, \dots, M$ ) in terms of the  $N(\mathbf{z}_j)$  ( $j = m + 1, \dots, M$ ) and  $N(\mathbf{0})$ . We define  $X_0(\mathbf{z}) \triangleq A(\mathbf{z})$ . In the  $m$ -th step of our (recursive) calculation, we assume that  $X_{m-1}(\mathbf{z}_m)$  has already been defined and that the following equation holds:

$$N(\mathbf{z}_m) = \frac{z_m}{z_m - S_m(X_{m-1}(\mathbf{z}_m))} \left\{ \frac{z_M X_{m-1}(\mathbf{z}_m) - S_M(X_{m-1}(\mathbf{z}_m))}{z_M} N(\mathbf{0}) + \sum_{j=m+1}^M \left[ \frac{S_j(X_{m-1}(\mathbf{z}_m))}{z_j} - \frac{S_{j-1}(X_{m-1}(\mathbf{z}_m))}{z_{j-1}} \right] N(\mathbf{z}_j) \right\}. \quad (3)$$

Substituting  $m = 1$  in this equation yields equation (2), which is the starting point of our recursive procedure. Applying Rouché's theorem, it can then be proved that for given values of  $z_j$  with  $|z_j| < 1$  ( $j = m + 1, \dots, M$ ), the equation  $z_m = S_m(X_{m-1}(\mathbf{z}_m))$  has a unique solution in the complex unit circle for  $z_m$ , which will be denoted by  $Y_m(\mathbf{z}_{m+1})$  in the remainder, and which is implicitly defined by  $Y_m(\mathbf{z}_{m+1}) \triangleq S_m(X_{m-1}(\mathbf{z}_m))|_{z_m=Y_m(\mathbf{z}_{m+1})}$ . Since any pgf is finite inside the unit circle and since  $Y_m(\mathbf{z}_{m+1})$  is a zero of the denominator of the right hand side of (3),  $Y_m(\mathbf{z}_{m+1})$  must also be a zero of the numerator. Defining  $X_m(\mathbf{z}_{m+1}) \triangleq X_{m-1}(\mathbf{z}_m)|_{z_m=Y_m(\mathbf{z}_{m+1})}$  (and  $X_0(\mathbf{z}_1) = A(\mathbf{z})$ ), this leads to expression (3) with  $m$  substituted by  $m + 1$ , which means that the  $m + 1$ -th step of the algorithm can be applied next. After  $M - 1$  iterations we finally find:

$$N(\mathbf{z}_M) = N(\mathbf{0}) \frac{z_M X_{M-1}(\mathbf{z}_M) - S_M(X_{M-1}(\mathbf{z}_M))}{z_M - S_M(X_{M-1}(\mathbf{z}_M))}. \quad (4)$$

Next, we can calculate the functions  $N(\mathbf{z}_m)$  ( $m = 1, \dots, M - 1$ ) as a function of  $N(\mathbf{0})$ . We therefore iteratively substitute the (in that step already) found expressions of  $N(\mathbf{z}_j)$  ( $j = m + 1, \dots, M$ ) in equation (3). Equating  $m$  to 1 finally gives  $N(\mathbf{z})$  as a function of  $N(\mathbf{0})$ . The expression for general  $M$  is too elaborate though, but we have outlined the principle by which this  $M$ -th dimensional function can be calculated. The last remaining unknown is  $N(\mathbf{0})$ . This constant can be calculated by applying the normalization condition  $N(\mathbf{1}) = 1$ , with  $\mathbf{1}$  a vector of size  $M$  with all elements equal to 1. This concludes the procedure to calculate  $N(\mathbf{z})$ , which is used in the analysis of the packet delays in the next section.

## 4 Packet Delays

The delay of a packet is defined as the number of slots between the end of the packet's arrival slot and the end of its departure slot. We denote the delay of a tagged class- $j$  packet by  $d_j$  and its pgf by  $D_j(z)$  ( $j = 1, \dots, M$ ). We furthermore denote the arrival slot of the tagged packet by slot  $k$ . If slot  $k$  is a start-slot, it is assumed to be start-slot  $l$ . If slot  $k$  is not a start-slot on the other hand, the last start-slot preceding slot  $k$  is assumed to be start-slot  $l$ . In this section, we show how an expression for  $D_j(z)$  - for general  $j$  - is derived.

Let us first refer to the packets in the system at the end of slot  $k$ , but that have to be served before the tagged packet as the "primary packets". So, basically, the tagged class- $j$  packet enters the server, when all primary packets and all packets with higher priority that arrived after slot  $k$  (i.e., while the tagged packet is waiting in the queue) are transmitted. In order to analyze the delay of the tagged class- $j$  packet, the number of packets that are served between the arrival slot of the tagged class- $j$  packet and its departure slot is important (and more specifically the time necessary to transmit them), not the precise order in which they are served. Therefore, in order to facilitate the analysis, we will consider an equivalent virtual system with an altered service discipline. From slot  $k$  on, we aggregate the  $j - 1$  highest priority classes in one class and serve

the packets in this aggregated class in a LCFS way (those in the queue at the end of slot  $k$  and newly arriving ones). So, a primary packet can enter the server, when the system becomes free (for the first time) of packets of this aggregated class that arrived during and after the service time of the primary packet that preceded it according to the new service discipline. Let  $v_{i,m}^{(n)}$  ( $i = 1, \dots, j$ ) denote the length of the time period during which the server is occupied by the  $m$ -th class- $i$  packet that arrives during slot  $n$  and its “successors” of the aggregated class, i.e., the time period starting at the beginning of the service of that packet and terminating when the system becomes free (for the first time) of packets of the  $j - 1$  highest priority classes which arrived during and after its service time. The  $v_{i,m}^{(n)}$ 's ( $i = 1, \dots, j$ ) are called sub-busy periods, initiated by the  $m$ -th class- $i$  packet that arrived during slot  $n$ . Notice that the  $v_{i,m}^{(n)}$  depend on the class we are analyzing (i.e. class- $j$ ), but to alleviate the notation this dependency is taken into account implicitly. It is clear that the lengths of consecutive sub-busy periods initiated by class- $i$  packets are i.i.d. and thus have the same pgf  $V_i(z)$  (which implicitly depends on  $j$ ). This pgf is given by

$$V_i(z) = S_i(zA(V_1(z), \dots, V_{j-1}(z), 1, \dots, 1)), \quad (5)$$

with  $i = 1, \dots, j$ ;  $j = 1, \dots, M$ . This can be understood as follows: when the  $m$ -th class- $i$  packet that arrived during slot  $n$  enters service, its sub-busy period,  $v_{i,m}^{(n)}$ , consists of two parts: the service time of that packet itself, and the service times of the packets of higher priority than the tagged class- $j$  packet (the aggregated class) that arrive during its service time and of their successors of the aggregated class. This leads to equation (5).

Finally,  $d_j$  can be expressed in terms of the  $n_{i,l}$ ,  $i = 1, \dots, M$  defined in the previous section.  $D_j(z)$  is then calculated as a function of  $N(\mathbf{z})$  by  $z$ -transforming this expression for  $d_j$ . The function  $N(\mathbf{z})$  was already calculated in section 3 and as a result  $D_j(z)$  can be found. We refer to [7] and [8] for more details on similar queueing analyses.  $D_j(z)$  is found to be given by (after some extensive mathematical manipulations)

$$D_j(z) = \frac{1 - \sum_{i=1}^j \rho_i S_j(z)(zB_{j-1}(z) - 1) B_j(z) - B_{j-1}(z)}{\lambda_j zB_{j-1}(z) - B_j(z) V_j(z) - 1} \quad (6)$$

$$\times \left( 1 - \frac{\sum_{i=j+1}^M \rho_i}{1 - \sum_{i=1}^j \rho_i} + \frac{1}{1 - \sum_{i=1}^j \rho_i} \sum_{i=j+1}^M \rho_i \frac{V_i(z) - 1}{\mu_i(zB_{j-1}(z) - 1)} \right),$$

with expression (5) of the  $V_i(z)$  expanded to  $i = j + 1, \dots, M$  and with  $B_i(z) \triangleq A(V_1(z), \dots, V_i(z), 1, \dots, 1)$  ( $i = 1, \dots, j$ ). Note that this expression is also correct for  $D_1(z)$ , the pgf of the highest priority class.

## 5 Performance Measures

The functions  $V_i(z)$ , defined in the previous section, can be explicitly found only in case of some specific arrival and service processes. Their derivatives for  $z = 1$ ,

necessary to calculate the moments of the packet delay, on the contrary, can be calculated in closed-form. So means, variances and higher moments of the packet delays of all classes can be calculated straightforwardly by taking the appropriate derivatives of expression (6) and substituting  $z$  by 1.

Furthermore, the tail probabilities of the packet delays can also be approximately calculated from the pgf's calculated in the previous section. These tail distributions are often used to impose statistical bounds on the guaranteed QoS for both classes. In order to determine the asymptotic behavior of the tail distribution, the dominant singularity of the respective pgf is important. The tail behavior of the delay of class- $j$  packets is a bit more involved than in usual queueing analyses, since it is not a priori clear what the dominant singularity of  $D_j(z)$  is. This is due to the occurrence of the functions  $V_i(z)$ ,  $i = 1, \dots, j - 1$  in (6), which are only implicitly defined. In [6] it is proved that these implicitly defined functions have a branch-point singularity  $z_B$  where their first derivatives become infinite but the functions themselves remain finite.  $z_B$  is then also a branch point of  $D_j(z)$ . A second potential singularity of  $D_j(z)$  is given by the (dominant) zero  $z_P$  of  $zB_{j-1}(z) - B_j(z)$  on the real axis (see expression (6)).

The tail behavior of the packet delay of class- $j$  packets is thus characterized by the singularities  $z_P$  or  $z_B$ , depending on which one is dominant (i.e., which one has the smallest modulus). Three types of tail behavior may occur, namely when  $z_P$  is dominant,  $z_P = z_B$  and  $z_B$  is dominant. In those three cases, the tail probabilities of the class- $j$  packet delay are given by (see [6] for more details)

$$\text{Prob}[d_j = n] \approx \begin{cases} K_1 z_P^{-n+1} & \text{if } z_P \text{ dominant} \\ \frac{K_2 n^{-1/2} z_B^{-n}}{\sqrt{z_B \pi}} & \text{if } z_P = z_B \text{ dominant} \\ \frac{K_3 n^{-3/2} z_B^{-n}}{2\sqrt{\pi/z_B}} & \text{if } z_B \text{ dominant.} \end{cases} \quad (7)$$

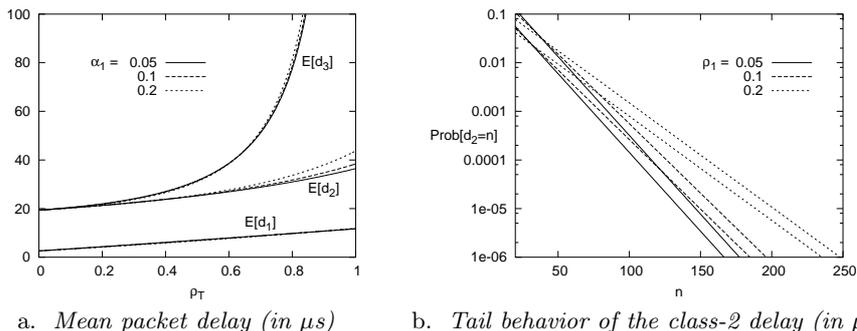
The constants  $K_i$  ( $i = 1, 2, 3$ ) can be found by investigating the pgf  $D_j(z)$  in the neighborhood of its singularity. The first expression of (7) shows a typical geometric tail behavior, the third expression shows a typical non-geometric tail behavior and the second expression gives a transition between both.

## 6 Numerical Examples

In this section, we present some numerical examples. We assume traffic of three traffic classes being handled by a queue, e.g. a class consisting of voice traffic, one of Video-on-Demand (VoD) traffic and a third one of data traffic. We call these class-1, class-2 and class-3 respectively in the remainder. Evidently, an interactive voice application will have the most stringent delay requirements while data (file transfer) will have the loosest delay bounds, with VoD somewhere in between. This is reflected by the priority level that has been assigned to each of the three traffic classes. The numbers of per-slot arrivals of class- $j$  are distributed according to a Poisson process with arrival rate  $\lambda_j$ . We furthermore assume

deterministic service times for class- $j$  equal to  $\mu_j$  number of slots. The arriving packets are transmitted via an output link. We assume that this link has a transmission rate of 620 Mb/s. The video and data packets all carry 1500 bytes corresponding to the length of an Ethernet packet. Due to the relatively low bitrate of a voice codec (8-64 kb/s), the filling time of a voice packet can become significant if the packet length is too high; as a result voice packets are usually kept small and are chosen to be equal to 200 bytes in this section. The slot-length  $\Delta$  then equals the amount of time to transmit 100 bytes, or  $\Delta = 1.29\mu s$ . Define  $\alpha_j$  ( $j = 1, 2, 3$ ) as the fraction of load of class- $j$  in the total traffic mix, i.e.,  $\alpha_j = \rho_j / \rho_T$ .

In Figure 1a., the means of the packet delay of the class-1, class-2 and class-3 packets are shown as functions of the total load  $\rho_T$  for  $\alpha_1 = 0.05, 0.1$  and  $0.2$  respectively and for  $\alpha_2 = 0.4$ . One can observe the influence of assigning priorities: mean delay of voice packets is kept as low as possible. Since the voice packets constitute a limited fraction of the traffic stream, the mean delay of the video packets is also kept relatively low. However, one can see that, especially for  $\alpha_1 = 0.2$  and for high loads, the influence of the voice packets on the mean delay of the video packets is not negligible. The price to pay for limiting the delay of voice and video packets is a larger mean delay of the data packets (as expected). This figure also shows that the mean delays of *all* classes suffers as the share of the high-priority traffic in the overall load increases.



**Fig. 1.** Numerical examples

Figure 1b. shows the tail probabilities of the packet delay of the class-2 (video) packets with  $\rho_1 = 0.05, 0.1$  and  $0.2$  respectively and  $\rho_2 = 0.4$ . The value of  $d_2$  of this figure is expressed in  $\mu s$ . For each value of  $\rho_1$  we have plotted two curves. The lower curves depict the tail probabilities when no data traffic arrives (i.e.,  $\rho_3 = 0$ ). The upper curves show the tail probabilities when the bandwidth that is not used by the voice or video traffic is consumed entirely by data packets ( $\rho_3 = 1 - \rho_1 - \rho_2$ ). We see that data packets have a non-negligible influence on the delay characteristics of video traffic (since the priority scheduling is non-preemptive and thus video packets that arrive during the transmission of a data

packet have to wait until that packet is fully transmitted), although this influence remains limited. The impact of the voice packets on the delay characteristics of the video packets is much larger. This was to be expected because of the priority given to voice packets over video packets.

## 7 Conclusions

In this paper, we analyzed the packet delays of all classes in a queueing system with a non-preemptive (HOL) priority scheduling discipline and with a general number of priority classes. A generating-functions-approach was adopted, which led to closed-form expressions of the moments and accurate approximations of the tail probabilities of the packet delays of all the classes, that are easy to evaluate. The service times are class-based and generally distributed. Therefore, the results could be used to evaluate the system performance in packet-based networks, that support multiple applications to which different priorities are assigned. An example is touched upon wherein voice, video and data streams are multiplexed.

## References

1. Rubin, I., Tsai, Z.: Message delay analysis of multiclass priority TDMA, FDMA, and discrete-time queueing systems. *IEEE Transactions on Information Theory* **35** (1989) 637–647
2. Takahashi, Y., Hashida, O.: Delay analysis of discrete-time priority queue with structured inputs. *Queueing Systems* **8** (1991) 149–164
3. Takine, T., Matsumoto, Y., Suda, T., Hasegawa, T.: Mean waiting times in non-preemptive priority queues with Markovian arrival and i.i.d. service processes. *Performance Evaluation* **20** (1994) 131–149
4. Sugahara, A., Takine, T., Takahashi, Y., Hasegawa, T.: Analysis of a nonpreemptive priority queue with SPP arrivals of high class. *Performance Evaluation* **21** (1995) 215–238
5. Takine, T.: A nonpreemptive priority MAP/G/1 queue with two classes of customers. *Journal of Operations Research Society of Japan* **39** (1996) 266–290
6. Laevens, K., Bruneel, H.: Discrete-time multiserver queues with priorities. *Performance Evaluation* **33** (1998) 249–275
7. Walraevens, J., Steyaert, B., Bruneel, H.: Delay characteristics in discrete-time GI-G-1 queues with non-preemptive priority queueing discipline. *Performance Evaluation* **50** (2002) 53–75
8. Bruneel, H., Kim, B.: Discrete-time models for communication systems including ATM. Kluwer Academic Publisher, Boston (1993)