Manuscript submitted to
AIMS' Journals
Volume **X**, Number **0X**, XX **200X**

Website: http://AIMsciences.org

pp. **X–XX**

FRAME-BOUND PRIORITY SCHEDULING IN DISCRETE-TIME QUEUEING SYSTEMS

Sofian De Clercq,

Koen De Turck, Bart Steyaert, and Herwig Bruneel

SMACS Research Group, Ghent University
St.-Pietersnieuwstraat 41,
9000 Gent, BELGIUM

Abstract. A well-known problem with priority policies is starvation of delay-tolerant traffic. Additionally, insufficient control over delay differentiation (which is needed for modern network applications) has incited the development of sophisticated scheduling disciplines. The priority policy we present here has the benefit of being open to rigorous analysis. We study a discrete-time queueing system with a single server and single queue, in which N types of customers enter pertaining to different priorities. A general i.i.d. arrival process is assumed and service times are generally distributed. We divide the time axis into 'frames' of fixed size (counted as a number of time-slots), and reorder the customers that enter the system during the same frame such that the high-priority customers are served first. This paper gives an analytic approach to studying such a system, and in particular focuses on the system content (meaning the customers of each type in the system at random slotmarks) in stationary regime, and the delay distribution of a random customer. Clearly, in such a system the frame's size is the key factor in the delay differentiation between the N priority classes. The numerical results at the end of this paper illustrate this observation.

1. **Introduction.** In modern packet-based communication networks the provisioning of adequate QoS (Quality of Service) guarantees to different traffic flows is often problematic. Different network applications demand different QoS from the network layer underneath. Commonly we can distinguish two types of traffic. One having higher delay-tolerance, such as e-mail, and VoD (video-on-demand) where jitter is more problematic. The other, being real-time applications, is less delay-tolerant. Here we think of VoIP, or video conferencing.

Answers to this demand include for instance weighted round robin (WRR), weighted fair queueing (WFQ) ([10]), and place reservation ([4]). The most extreme way of priority scheduling is absolute priority (AP) or HoL-priority (Head of Line), either preemptive or non-preemptive (see f.i. [14]). In this scheduling discipline,

2000 *Mathematics Subject Classification.* Primary: 60K25, 68M20; Secondary: 90B22. This paper was presented at the QTNA2010 conference which was held in Beijing, China, during July 24-26, 2010. An earlier and brief version of this paper was published in the Conference Proceedings. The reviewing process of the paper was handled by Wuyi Yue and Yutaka Takahashi as Guest Editors.

Key words and phrases. Queueing theory, priority, discrete-time.

the highest priority packets enter the queue at the head of the line. This creates the problem of starvation, where delay-tolerant traffic can suffer long waiting periods as high priority packets keep skipping to the head of the line, when traffic loads are high.

In this paper we present a solution to the starvation problem that can be tuned to offer high delay differentiation between traffic classes. The idea is to divide the time axis in time-frames and let a high-priority packet overtake low-priority packets if they have entered the buffer (hereafter the queue) during the same time-frame. This way the sojourn time of a low-priority packet will be fixed once the frame during which it entered finishes, hence cancelling starvation. We will present a method to analyse the system content in such a system using matrix analytical tools (as in [1]) as well as probability generating functions ([15]). Grouping packets that entered the buffer during such a frame (much like the groups introduced in [12]) will prove invaluable in obtaining analytic results for this particular service mechanism. From now on, we will refer to 'packets' as customers.

In section 2 we propose a discrete-time queueing model that allows delay differentiation among multiple customer classes. We will introduce some notations that will be used throughout the paper. Next follows the analysis of the system, in which we aim for the joint pgf of the system content at random slot marks, after which we briefly touch upon the different delay distribution's pgfs (for the different customer classes). Finally some numerical results show the effect of the frame size on the delay differentiation between customers of distinct types.

2. Model and Definitions.

We consider a discrete-time queueing system in which customers of N types arrive according to a general i.i.d. arrival process. Let $a_{j,n}$ be the number of type-j customers entering the system during slot n. Then the joint probability generating function (pgf) $E[\prod_{j=1}^{N} z_j^{a_{j,n}}]$ is independent of the slot-index n, and we abbreviate it as $A(\mathbf{z})$ with $\mathbf{z}$ the vector with j-th entry z_j. Service times of customers of type j are generally distributed and i.i.d. random variables we denote by s_j (with pgf $S_j(z) \triangleq E[z^{s_j}]$), and service of a customer can start no sooner than the slot following its arrival in the system - even when there are no customers being served during this slot. As our notation implies, service times of customers of type-i and type-j $(i \neq j)$, can have different distributions and are independent of one another.

The adopted scheduling mechanism can be better understood when compared against the FHLL (First High, Last Low priority) principle ([13]) or slot-bound priority ([3]), in which a tagged type-t customer has priority over all customers of type-u that entered the system during the same slot as our tagged customer $(\forall u > t)$. Customers entering the system during different slots are served FCFS (first-come-first-served) regardless of their type. In this scheduling mechanism, priority is limited to customers entering the system during the same slot. Its effect appears to be quite limited for most traffic scenarios that were considered in [3]. The idea to generalize the priority to take effect on customers entering the system during fixed-size intervals (called frames hereafter) follows quite naturally from this observation. The actual analysis on the other hand does not, since several additional complications are to be taken into account, as we will show in the remainder of this paper. Moreover note that because in [13] the FHLL principle is applied to a finite (multi-threshold) queueing system, and studied using network calculus, it offers little basis to obtain the pgfs of delay and queue content in our model.

We divide the time axis in frames of equal size (each equal to M slots). Frame-bound priority is defined such that customers entering the system during the same frame are served according to their priority class (1 being the highest priority and are thus served first, N being the lowest priority class). Customers entering the system during different frames are to be served FCFS. This poses a first problem. Namely, suppose that at some time instant all customers that entered the system during previous frames have been served. A customer that enters the system during the running frame can then be served immediately. But if after that customer, a customer of a higher priority class arrives during the same frame, then the service order would be reversed among customers with arrival instances during the same frame. To counter this, a possible solution could be to delay service of a customer entering the system while it is empty until the running frame terminates. As this leaves the system idle while there are customers present in the system, the work conservation principle would be violated, and we propose another solution: if at some point in time all customers that entered during past frames have been served before the length of the running frame has reached M, the running frame is terminated and the service of the customers that belong to it (if any) commences, according to their respective priority level. In this way the prioritized service order within frames is guaranteed. Frames however may have lengths of less than M slots due to this principle, and in particular all frames during idle periods are of length 1 (see Fig.1).

To aid our analysis, we keep track of the ordinal number of a slot within a frame (ranging from 1 to M, M being the maximum frame length). We refer to the ordinal number itself as the phase. The first slot in each frame will have phase 1. We say that the current frame is reset or terminated when all customers that entered before the most recent frame bound have been served.

We define a group (of order l, $l \in \{1, \dots, M\}$) as the collection of customers that entered the system during the same frame, given that the frame's length is l slots. As our discussion shows, groups are served FCFS, and between customers of the same group an absolute priority rule holds, as in [3]. Here however we have the additional difficulty that we have groups of different orders and thus no identically distributed (group) service times, but depending on the system state at the time of service initiation (where with a groups service time we mean the combined service time of all customers that are part of it). Hence the basic steps used in the analysis of the slot-bound priority in [3] cannot be applied here.

Say that during a frame of length l slots no customers enter the system, then the formed group is empty. Since however such groups will possibly have zero service time while customers have non-zero service time, some complications may arise. Therefore we will only consider groups when at least one customer enters the system during tagged frame. Let $a_{g,k}^{(l)}$ be the indicator for a group entering the system during the k'th frame when this frame's size is l slots (a_g stands for arrival of groups). Because of the i.i.d. property of the arrival process we can see that these indicators are independent of k, and their pgf is given by

$$A_g^{(l)}(z) \triangleq E[z^{a_g^{(l)}}] = A(\mathbf{0})^l + (1 - A(\mathbf{0})^l)z, \tag{1}$$

where we have omitted the frame index k. The symbols $\mathbf{0}$ and later $\mathbf{1}$ are vectors of which all entries are equal to 0 and 1 respectively. Furthermore, the number of type-j customers in a group of order l is a discrete random variable (drv) which we denote by $b_j^{(l)}$. It is basic to see that $b_j^{(l)}$'s distribution is the same as $\sum_{n=1}^{l} a_{j,n}$'s conditioned on the subspace where during the entire l slots no customers of any type enter the system ($\sum_{n=1}^{l}\sum_{j=1}^{N} a_{j,n} > 0$). The joint pgf of $b_1^{(l)}, \ldots, b_N^{(l)}$ hence follows from the above, with the result

$$B^{(l)}(\mathbf{z}) \triangleq E[\prod_{j=1}^{N} z_j^{b_j^{(l)}}] = \frac{A(\mathbf{z})^l - A(\mathbf{0})^l}{1 - A(\mathbf{0})^l}. \tag{2}$$

All these definitions are illustrated in Fig.1. We also define $a_{T,n}$ (with pgf $A_T(z) = A(z, z, \ldots, z)$) as the total number of customers entering the system during slot n ($= \sum_{j=1}^{N} a_{j,n}$), and its average as $\lambda_T \triangleq E[a_T]$, which is again independent of n. Likewise, we write $\lambda_j \triangleq E[a_j]$ for short. Since the system we described is work-conserving, we deduce that the fraction of time during which the server serves a type-j customer is given by $\rho_j \triangleq \lambda_j E[s_j]$. In the analysis we will assume that the system evolves to a stationary regime, i.e. $\sum_{j=1}^{N} \rho_j \triangleq \rho < 1$.

We aim to obtain the joint pgf of v_j, ($V(\mathbf{z}) \triangleq E[\prod_{j=1}^{N} z_j^{v_j}]$): the number of type-$j$ customers in the system at random slot marks, for all $j \in \{1, \ldots, N\}$, assuming our system reaches a steady-state.

Before focussing on the system content analysis, we want to derive some preliminary results concerning the group-forming procedure. One can view groups as a sort of super-customer. Their service times are i.i.d. with the exception of the group at the head of the line - when the most recent group departure caused the running frame to reset (a group leaves the system after its last customer is served) - which is not necessarily of order M. We agree that a group arrival is perceived to take place during the last slot of the frame it originates from, and likewise we say that a group enters the system at such an epoch.

Consider a period in time during which the frames all have their maximum length of M slots. We wish to know the probability that a group will enter the system during a slot, given the phase during that slot. Next to that we wish to keep track of the phase in the next slot. Both questions are answered by the following matrix

$$X(z) \triangleq \begin{pmatrix} & 1 & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \\ A_g^{(M)}(z) & & & & \end{pmatrix} . \tag{3}$$

The above is a probability generating matrix and $(X(z)\mathbf{1})_l$ - '$\mathbf{1}$' is a $M \times 1$ column vector with all entries equal to 1, not to be confused with the parameterlist when used as a multivariate functions argument - is the pgf of the number of group arrivals (one or none) during the tagged slot given the phase is l. We can see that all nonzero entries in this column vector are equal to 1 except for the M'th which equals $A_g^{(M)}(z)$. This guarantees that group arrivals occur during the last slot of the frame of length M only. Furthermore $(X(1))_{i,j}$ is the probability that the phase in the next slot will be j given that the phase in the current one was i. $X(z)$ guarantees that we go from phase i to phase $i+1$ in a deterministic manner during successive slots in a frame of length M. Also note that $(X(z)^n\mathbf{1})_l$ is a row vector that describes the number of group arrivals during n consecutive slots given the starting phase was l, and $(X(1)^n)_{i,j}$ represents the probability that the phase will evolve to i during these n slots given it started in phase j.

Let us denote the combined service time of all customers making up a random group of order l as $s_g^{(l)}$. If we are again interested in the number of group arrivals during this group's service time and the phase we end in, given that we start in phase i, we can utilize the previously introduced $X(z)$ since we know that no frames are cut short during such a service time with the exception of maybe its last frame. We will neglect this possibility for reasons that will become clear in the analysis itself. The answer can be formulated in the same way as above, this time with the matrix $S_g^{(l)}(X(z))$, in which the l stands for the order of the tagged group.

$$s_g^{(l)} \triangleq \sum_{j=1}^{N} \sum_{i=1}^{b_j^{(l)}} s_{j,i}, \tag{4}$$

$$S_g^{(l)}(z) \triangleq E[z^{s_g^{(l)}}] = B^{(l)}(\mathbf{S}(z)), \tag{5}$$

$$\begin{aligned} S_g^{(l)}(X(z)) &\triangleq E[X(z)^{s_g^{(l)}}] \\ &= \sum_{i=1}^{\infty} X(z)^i \Pr[s_g^{(l)} = i] \, , \end{aligned} \tag{6}$$

where $\mathbf{S}(z)$ represents an N-dimensional vector with j-th entry equal to $S_j(z)$, pgf of the service time of an arbitrary type-j customer. The last definition is that of a matrix generating function with comparable properties to $X(z)$, namely that $(S_g^{(l)}(X(1)))_{i,j}$ is the probability that starting from phase i the phase will evolve to phase j once the service time of a group of order l has ended. Furthermore

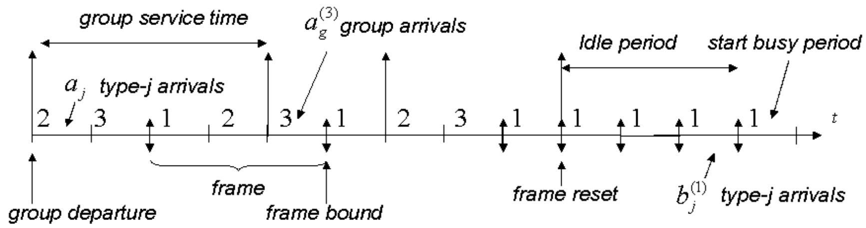

FIGURE 1. Time axis with frame bounds and group service times. The phase is indicated above each slot and $M = 3$.

$(S_g^{(l)}(X(z))\mathbf{1})_j$ is the pgf of the number of group arrivals during this service time given that the phase started out in j. We will use these functions elaborately in the following analysis.

3. **Analysis.**

In our analysis, we will adopt the indices n and k to refer to slot and frame (or group) numbers respectively.

3.1. **A Renewal-Points Approach.**

We say that a group leaves the system when its last customer has been served. Likewise, with a group departure epoch, we refer to the observation point of such an event. Immediately after a group departure epoch all groups will have the property that none of their customers will have been served. However, defining the system state as the number of groups in the system and the phase of the slot right after a group departure epoch does not constitute a Markov chain. The problem occurs when both drv's are 1, in which case one cannot be sure of the order of the only group in the system. Knowledge of the order of this group is necessary in determining the number of slots until the next group departure. Furthermore, previous states can help in obtaining the order of this group, and so the Markov property is not fulfilled.

In the remainder of the analysis we will therefore denote the number of groups in the system at the first slot mark *preceding* the k'th group departure (and thus including the departing group) by u_k and the phase during the last slot of this groups service time as m_k. Naturally $u_k > 0$, $\forall k$ because a group's service time, as it contains at least one customer, is strictly positive and thus u_k includes the group being served that leaves the system one slot thereafter. Because our state space is a semi-infinite strip, (m_k can only be one of M different values), we will opt to obtain a vector pgf for the steady-state distribution of (u_k, m_k). We define the row-vector

$$\begin{aligned}(\mathbf{U}(z))_l &\triangleq E[z^u|m=l]\Pr[m=l], \ \ \forall l \in \{1,\dots,M\} \\ &= \sum_{i=0}^{\infty} z^i \Pr[u=i, m=l], \ \ \forall l \in \{1,\dots,M\} \,. \end{aligned} \tag{7}$$

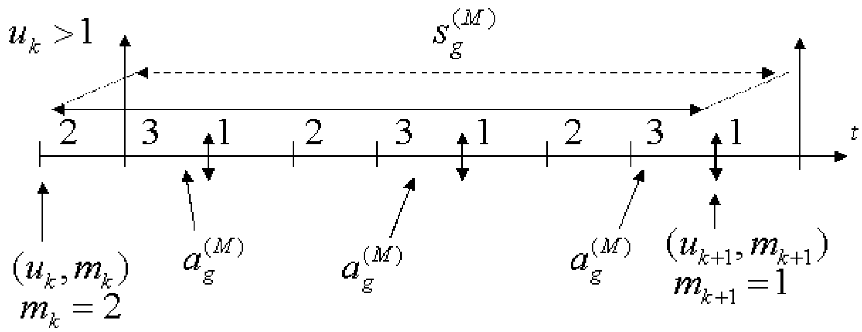

FIGURE 2. If $u_k > 1$ the next group to be served will be of order M. Hence, the two observation epochs are exactly one service time of a group of order M apart ($M = 3$).

Since $u > 0$, τ_l being the boundary vector's l'th entry will be given by $\Pr[u = 1, m = l]$. We will denote this boundary vector by $\boldsymbol{\tau}$. In the following we will derive a closed form formula for $\mathbf{U}(z)$ as a function of the arrival process $A(\mathbf{z})$ and the service time distributions in $S_j(z)$.

The adopted strategy is to relate (u_k, m_k) to (u_{k+1}, m_{k+1}). In doing so we distinguish three distinct cases. A first being $u_k > 1$, where we know that the next group to be served following the group departure this k'th observation epoch is associated to (we will later refer to it as group K) is of order M. This first case covers the non-linear terms in $\mathbf{U}(z)$ being $\mathbf{U}(z) - \boldsymbol{\tau} z$. The second and third case both have $u_k = 1$, where we distinguish whether during the frame that is being reset customers enter or not. In the former case a group of order m_k is served starting right after the tagged group departure. In the latter case an idle period starts after the relevant group departure until a batch of customers enters the system resulting in a group of order 1.

So starting with the first case ($u_k > 1$), we can see that u_{k+1} is made up of all groups in u_k except for the group that left one slot after u_k was probed. On top of that, it counts all groups that entered the system during the service time of group K (see Fig.2) excluding the last slot before its departure but including the slot before its service initiation. The period between the k'th and $(k+1)$'th observation epochs is exactly a service time of a group of order M, and no frames are reset in between. We argued that the matrix generating function of the number of group arrivals in this period was given by $S_g^{(M)}(X(z))$ conditioned on the starting phase, being m_k, and thus

$$\left[(\mathbf{U}(z)z^{-1} - \boldsymbol{\tau})S_g^{(M)}(X(z))\right]_j = \sum_{i=1}^{\infty} z^i \Pr[u_{k+1} = i, m_{k+1} = j, u_k > 1] \,, \qquad (8)$$

Next, when $u_k = 1$ the running frame is reset at the group departure epoch. Two cases flow forth from this. In the first case depicted in Fig.3, there where no

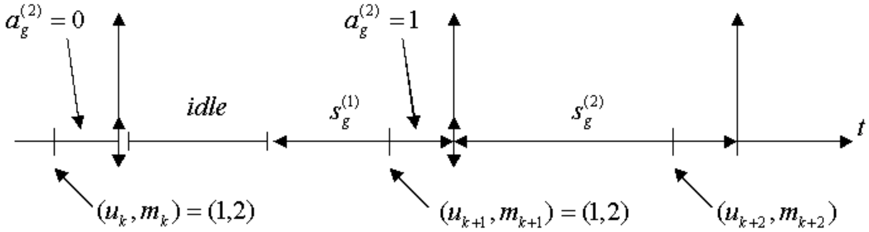

FIGURE 3. When $u_k = 1$, we distinguish two cases. One where $a_g^{(m_k)} = 0$ and one where it is 1. Each has its own effect on the order of the group that is next to be served.

customer arrivals during the incomplete frame which lasted m_k slots. An idle period follows this group departure, which is ended by the arrival of a group of customers initiating a busy period. This busy period starts of by the service of this group of order one during the following $s_g^{(1)}$ slots. Because the $(k+1)$'th observation epoch is one slot prior to this group's departure whose service started in phase one, we know that the vector pgf of the number of group arrivals and the ending phase, is given by $\left(S_g^{(1)}(X(z))X(z)^{-1}\right)_{1*}$. For u_{k+1} this yields

$$\Big[\sum_{l=1}^{M}\tau_l A_g^{(l)}(0) z\left(S_g^{(1)}(X(z))X(z)^{-1}\right)_{1*}\Big]_j = \sum_{i=1}^{\infty} z^i \times \sum_{l=1}^{M}\Pr[u_{k+1}=i, m_{k+1}=j, u_k=1, m_k=l, a_g^{(l)}=0]. \tag{9}$$

The index $(.)_{1*}$ is used to represent the first row of the matrix argument - the service of the departing group starts in phase one.

If $u_k = 1$ but this time a group does enter the system during the incomplete frame (which occurs with probability $1 - A_g^{(m_k)}(0)$), the group being served next will be of order m_k. Again, its service starts in phase one, and because the next observation epoch is one slot prior to the tagged group's departure, the vector pgf of the number of group arrivals and the ending phase is given by $\left(S_g^{(m_k)}(X(z))X(z)^{-1}\right)_{1*}$. For u_{k+1}, this results in

$$\Big[\sum_{l=1}^{M}\tau_l (1 - A_g^{(l)}(0)) z\left(S_g^{(l)}(X(z))X(z)^{-1}\right)_{1*}\Big]_j = \sum_{i=1}^{\infty} z^i \times \sum_{l=1}^{M}\Pr[u_{k+1}=i, m_{k+1}=j, u_k=1, m_k=l, a_g^{(l)}=1], \tag{10}$$

The result we were looking for (being an expression for $\mathbf{U}(z)$), is simply the sum of (8), (10), and (9), yielding

$$\begin{aligned}\mathbf{U}(z) =& (\mathbf{U}(z)z^{-1} - \boldsymbol{\tau})S_g^{(M)}(X(z)) \\ &+ \sum_{l=1}^{M} \tau_l z \left(S_g^{(l)}(X(z))\right)_{1*} X(z)^{-1}(1 - A_g^{(l)}(0)) \\ &+ \sum_{l=1}^{M} \tau_l z \left(S_g^{(1)}(X(z))\right)_{1*} X(z)^{-1} A_g^{(l)}(0) \ . \end{aligned} \tag{11}$$

The only unknown in this expression is the boundary vector $\boldsymbol{\tau}$ (see later). We can rewrite the above equation in the following form:

$$\mathbf{U}(z)(zI - S_g^{(M)}(X(z))) = \boldsymbol{\tau} z(zD(z) - S_g^{(M)}(X(z))) \ , \tag{12}$$

in which $D(z)$ is an $M \times M$ matrix with the l'th row given by

$$\begin{aligned}(D(z))_{l*} \triangleq& (1 - A_g^{(l)}(0)) \left(S_g^{(l)}(X(z))\right)_{1*} X(z)^{-1} \\ &+ A_g^{(l)}(0) \left(S_g^{(1)}(X(z))\right)_{1*} X(z)^{-1} \ . \end{aligned} \tag{13}$$

3.2. **Computation of the Boundary Vector.** A common strategy is to use the normalization condition to obtain the boundary vector, or in this case at least part of it. Deriving both sides of equation (12) with respect to z, evaluating in $z = 1$, and right-multiplying by $\mathbf{1}$, the $M \times 1$ column vector with all entries equal to one, yields:

$$1 - \rho = \boldsymbol{\tau} \left(\mathbf{1} + \frac{d}{dz} D(z) \Big|_{z=1} \mathbf{1} - \frac{d}{dz} S_g^{(M)}(X(z)) \Big|_{z=1} \mathbf{1} \right) \ . \tag{14}$$

Now we still need $M - 1$ additional equations to determine $\boldsymbol{\tau}$. In the following we outline a well-known matrix-analytic procedure to produce such equations. We can write $S_g^{(M)}(X(z))$ as its series expansion and we call the coefficient matrix of z^i, E_i. Instead of z we can adopt a matrix Z as argument as follows (and likewise with $D(z)$ with coefficient matrices F_i) :

$$S_g^{(M)}(X(Z)) = \sum_{i=0}^{\infty} E_i Z^i \tag{15}$$

$$D(Z) = \sum_{i=0}^{\infty} F_i Z^i \ . \tag{16}$$

Naturally these infinite sums converge only if all eigenvalues of Z lie in or on the unit disk. Observe that

$$(E_i)_{l,l'} = \Pr[u_{k+1} = u_k - 1 + i, m_{k+1} = l' | u_k > 1, m_k = l] \tag{17}$$

$$(F_i)_{l,l'} = \Pr[u_{k+1} = i, m_{k+1} = l' | u_k = 1, m_k = l] \ . \tag{18}$$

Obtaining the minimal nonnegative solution to $G = S_g^{(M)}(X(G))$, (whereby G is the fundamental matrix of the Markov chain) is the first step in solving for the boundary vector in the Neuts method (see f.i. [11]). From what we have found E_i to represent and the definition of G, we find that the (i,j)'th element of this G matrix is the probability that given a start phase of i, the first time u becomes lower than its initial value, it will end up in phase j. From this it follows that G is a stochastic matrix i.e. $G\mathbf{1} = \mathbf{1}$, and as is well known, the eigenvalues of a stochastic matrix all lie in or on the unit disk (see f.i. [6]). Say that we are interested only in the censored Markov chain restricted to level 0, or in our case $u = 1$. Then we obtain $\boldsymbol{\tau}$ as a result of the system of equations

$$\boldsymbol{\tau} D(G) = \boldsymbol{\tau} \ . \tag{19}$$

Given that we start from $u = 1$ and $m = i$, then the probability that the phase will be j during our next visit to $u = 1$, is given by $(D(G))_{i,j}$. Therefore the above equation represents the condition for having a steady-state solution for the boundary vector $\boldsymbol{\tau}$. Since $I - D(G)$ is singular we obtain not M but $M-1$ equations, which we can solve for $\boldsymbol{\tau}$ once combined with (14).

Usually the fundamental matrix is obtained in an iterative procedure with $S_g^{(M)}(X(.))$ as its kernel. This works because of the convergence of the series guaranteed in [1]. Computationally this can be quite heavy though but thanks to the cyclic nature of the $X(z)$ matrix, the complexity of solving $G = S_g^{(M)}(X(G))$, can be reduced significantly. Since $X(z)$ is linear in z, we rewrite $X(z)$ as $X_0 + X_1 z$, and the structure of $X(G)$ becomes apparent

$$X_0 \triangleq X(0), \qquad\qquad X_1 \triangleq X(1) - X(0)$$

$$X(G) = X_0 + X_1 G = \begin{pmatrix} & 1 & & \\ & & \ddots & \\ & & & 1 \\ X(G)_{M1} & X(G)_{M2} & \dots & X(G)_{MM} \end{pmatrix} . \qquad (20)$$

The above structure shows that $X(G)$ is a companion matrix. We denote the eigenvalues of $X(G)$ by ζ_k, and invoke a well-known property of companion matrices: the column eigenvector pertaining to eigenvalue ζ_k, is given by $(1, \zeta_k, \zeta_k^2, \dots, \zeta_k^{M-1})^T$. In the following $\boldsymbol{\nu}$ is an arbitrary column eigenvector of $X(G)$ pertaining to the eigenvalue ζ. If we apply $X(.)$ to the fundamental equation $G = S_g^{(M)}(X(G))$, we obtain from (20)

$$X(G) = X_0 + X_1 S_g^{(M)}(X(G)) = X_0 + X_1 \sum_{n \geq 0} s_g^{(M)}(n) X(G)^n \qquad (21)$$

in which $s_g^{(M)}(i) = \Pr[s_g^{(M)} = i]$ is a short-hand notation. Right multiplying each side by $\boldsymbol{\nu}$ yields a system of equations of which the first $M-1$ are trivial ones, but the last carries the information.

$$X(G)\boldsymbol{\nu} = \zeta\boldsymbol{\nu} = X_0\boldsymbol{\nu} + X_1 \sum_{n \geq 0} s_g^{(M)}(n)\zeta^n \boldsymbol{\nu} = X_0\boldsymbol{\nu} + X_1 \boldsymbol{\nu} S_g^{(M)}(\zeta) \qquad (22)$$

$$\zeta\boldsymbol{\nu} = \begin{pmatrix} \zeta^1 \\ \vdots \\ \zeta^{M-1} \\ \zeta^M \end{pmatrix} = \begin{pmatrix} \zeta^1 \\ \vdots \\ \zeta^{M-1} \\ A(\mathbf{0})^M \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ (1 - A(\mathbf{0})^M) S_g^{(M)}(\zeta) \end{pmatrix} . \qquad (23)$$

Using the definition of $A_g^{(M)}(z)$ and $S_g^{(M)}(z)$ we arrive at

$$\zeta^M = A_g^{(M)}(S_g^{(M)}(\zeta)) = A(\mathbf{S}(\zeta))^M \quad \Rightarrow \quad \zeta_k = A(\mathbf{S}(\zeta_k))\epsilon_M^k, \;\; k \in \{1, \dots, M\} \qquad (24)$$

in which ϵ_M represents an M'th primitive root of unity. From (24) it follows that the eigenvalues of $X(G)$ are the M solutions to $z^M = A(\mathbf{S}(z))^M$ inside the closed complex unit disk. Solving for these eigenvalues we find $X(G)$ to be given by

$$\begin{aligned} (V)_{ij} &= \zeta_j^{i-1}, \\ X(G) &= V \;\operatorname{diag}(\zeta_k)\; V^{-1}, \end{aligned} \qquad (25)$$

in which $\text{diag}(a_k)$, is an $M \times M$ diagonal matrix with k'th element a_k. Although G is easily calculated from $X(G)$, note that our main interest $D(G)$, can be obtained directly from $X(G)$ (notice that only $X(z)$ appears in (13)). For that purpose, we first write $D(z)$ in the following form (see (13)):

$$S_g^{(M)}(X(z)) = \frac{A(\mathbf{S}(X(z)))^l - A(\mathbf{0})^l I}{1 - A(\mathbf{0})^l}, \tag{26}$$

$$(E_{l1})_{ij} \triangleq \delta_{li}\delta_{1j}, \tag{27}$$

$$D(z) = \sum_{l=1}^{M} E_{l1}\left(A(\mathbf{S}(X(z)))^l + A(\mathbf{0})^l \frac{A(\mathbf{S}(X(z))) - I}{1 - A(\mathbf{0})}\right) X(z)^{-1}, \tag{28}$$

in which δ_{ij} is Kronecker's delta function. E_{l1} performs an elementary row operation which selects the first row of the matrix to the right of it and puts it on the l'th, much like in the definition of $D(z)$ (the other rows are zero). Using (24) and (25) we thus finally obtain

$$D(G) = \sum_{l=1}^{M} E_{l1} V \,\text{diag}\left(\zeta_k^{l-1}\epsilon_M^{-kl} + \frac{A(\mathbf{0})^l}{1 - A(\mathbf{0})}(\epsilon_M^{-k} - \zeta_k^{-1})\right) V^{-1} \tag{29}$$

The matrix equation $\boldsymbol{\tau} D(G) = \boldsymbol{\tau}$ can now be reduced to a set of equations for τ_l. In explicit terms, this yields

$$\sum_{l=1}^{M} \tau_l \frac{A(\mathbf{0})^l}{1 - A(\mathbf{0})}(\zeta_k \epsilon_M^{-k} - 1) = \sum_{l=1}^{M} \tau_l \zeta_k^l (1 - \epsilon_M^{-kl}) \;, \quad \forall k \in \{1, \ldots, M\} \;. \tag{30}$$

The case where $k = M$ is the trivial equation in this system, and confirms our previous statement, namely that $I - D(G)$ was in fact singular. Again, the normalization condition will form be the last equation needed to solve for $\boldsymbol{\tau}$. Concretely, the only difficulty remaining is solving the implicit equations (24), which can happen numerically, and is less time consuming than the iterative procedure for G.

3.3. **Random Slot Boundaries.** Now recall that our original interest was in $V(\mathbf{z})$, joint pgf of the number of customers of all types in the system at the beginning of a random slot. Because we already know from the definition of ρ that $V(\mathbf{0}) = 1 - \rho$ we choose a random slot during a busy period and call that slot, slot I. We define $(\hat{u}, \hat{m})$ as the number of groups in the system and the phase of the system, one slot before the most recent group departure since slot I, much like we did with (u, m) before, only this time we choose a random slot, as opposed to a random group departure. Let the drv T represent the type of the customer being served during

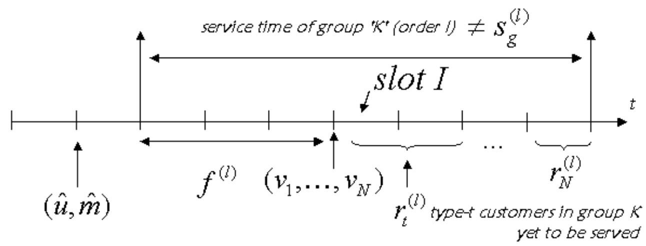

FIGURE 4. Definitions of $\hat{u}$, $\hat{m}$, $f^{(l)}$, and all $r_j^{(l)}$. Note that group K's service time is not given by $s_g^{(l)}$ as explained in the paragraph following equation (39).

slot I. We can find the following relation between (u, m) and $(\hat{u}, \hat{m})$.

$$\Pr[\hat{u} = i, \hat{m} = j | T = t, \hat{u} > 1] = \Pr[u = i, m = j | u > 1] \ . \tag{31}$$

Notice that the left hand side is independent of T. Let us denote the group that is served during slot I by group K. When $\hat{u} > 1$, then group K is of order M. Whatever $\hat{u}$ and $\hat{m}$'s values are has no further effect on group K's customer composition under this condition. Hence choosing a random slot under $\hat{u} > 1$ is like choosing a random group of order M with the previous group leaving the system with $u > 1$. So clearly T has no effect on the probability above. When $\hat{u} = 1$ the above equivalence does not hold, precisely because group K can be of order lower than M. This does not pose a problem since we will link it directly to the boundary vector $\boldsymbol{\tau}$ later on.

In this last case we will again have to distinguish between the cases where the last frame contained a group arrival, and where it did not (in which case an idle period starts and group K is of order 1). Let $f^{(l)}$ be the number of slots between the service initiation of group K and the beginning of slot I, given that group K is of order l. And let $r_j^{(l)}$ be the number of type-j customers in group K (order l), yet to be served at the beginning of slot I (including the type-t customer we conditioned on). For $\hat{u} > 1$ we can see from Fig.4 that

$$v_j = \sum_{i=1}^{\hat{u}-2} b_{j,i}^{(M)} + \sum_{n=1}^{f^{(M)}+\hat{m}} a_{j,n} + r_j^{(M)} \ , \quad \hat{u} > 1 \ . \tag{32}$$

The first sum goes over all groups in the system except for group K and the group that left the system before group K's service initiation but included in $\hat{u}$. It counts all type j customers in these groups and we have added an index i to $b_j^{(M)}$ to indicate that all drv's in this sum are i.i.d.. The type-j customers in group K that are not yet served at the beginning of slot I are accounted for in the third term of the above formula. Since the $\hat{u}$ groups were all formed at least $\hat{m}$ slots before group K's service initiation, we must not neglect to include the type-j customers that entered the system from that point on up to the beginning of slot I, spanning a full $f^{(M)}+\hat{m}$ slots. The second sum goes over all these slots.

In the case that $\hat{u}=1$ and the last frame counted at least one customer arrival (with probability $1-A(\mathbf{0})^{\hat{m}}$), then group K will be of order $\hat{m}$. The same train of thought can be followed as above, so the obtained formula for v_j is the same with the exception that M must be replaced by $\hat{m}$, and the $\hat{m}$ in the upper limit of the second sum must be omitted because arriving customers during the slots that this $\hat{m}$ represents, are part of group K and are counted by $r_j^{(\hat{m})}$. Notice that the first sum disappears in this case resulting in

$$v_j=\sum_{n=1}^{f^{(\hat{m})}} a_{j,n}+r_j^{(\hat{m})}\ ,\quad \hat{u}=1, a_g^{(\hat{m})}=1\ , \tag{33}$$

Lastly we have the case where, $\hat{u}=1$ and the last frame harbors no customer-arrivals (with probability $A(\mathbf{0})^{\hat{m}}$). Again the same principles can be used but this time group K is of order 1, and M should be replaced in (32) accordingly ($M\rightarrow 1$). Furthermore the $\hat{m}$ in the upper limit of the second sum must be omitted for the same reason as before. These observations lead to the result

$$v_j=\sum_{n=1}^{f^{(1)}} a_{j,n}+r_j^{(1)}\ ,\quad \hat{u}=1, a_g^{(\hat{m})}=0\ . \tag{34}$$

In (32), (33) and (34) all drv's are independent of one another except for $f^{(M)}$ and $r_j^{(M)}$ in (32), $f^{(l)}$ and $r_j^{(l)}$ in (33), and $f^{(1)}$ and $r_j^{(1)}$ in (34), who are correlated. For general l, $f^{(l)}$ is inversely proportional to the remaining service time of group K. On the other hand $r_j^{(l)}$ is proportional to the remaining service time. In the end we are interested in $\Pr[f^{(l)}=n, r_1^{(l)}=i_1,\ldots,r_N^{(l)}=i_N]$, but it will prove easier and more helpful to determine their joint partial pgf $H_t^{(l)}(\mathbf{x},z)$ (partial on t, and $\mathbf{x}$ being an N-dimensional vector with j'th element equal to x_j).

$$H_t^{(l)}(\mathbf{x},z)\triangleq\sum_{n=0}^{\infty} z^n\sum_{i_1=0}^{\infty} x_1^{i_1}\cdots\sum_{i_N=0}^{\infty} x_N^{i_N}\Pr[f^{(l)}=n, r_1^{(l)}=i_1,\ldots,r_N^{(l)}=i_N, T=t]\ . \tag{35}$$

It will be easier to find a closed form expression for this sum using the Snake Oil Method ([15]). We call a_j^* the number of type-j customers originally in group K, and $s_{j,i}^*$ the service time of the i-th type-j customer in group K. We can thus rewrite the unknown probability in the previous expression as follows

$$\sum_{k_1,\ldots,k_N=0}^{\infty} \sum_{j_{1,1},\ldots,j_{N,k_N}=1}^{\infty} \Pr[f^{(l)} = n, r_1^{(l)} = i_1, \ldots, r_N^{(l)} = i_N,$$
$$a_1^* = k_1, \ldots, a_N^* = k_N, s_{1,1}^* = j_{1,1}, \ldots, s_{N,k_N}^* = j_{N,k_N}, T = t] \,. \tag{36}$$

There's quite a few things we know about this probability, which in the remainder we will abbreviate $\Pr[f^{(l)}, \mathbf{R}, \mathbf{A}^*, \mathbf{S}^*, T = t]$ for practical reasons. $\mathbf{R}$ will then stand for all events of the form $r_p^{(l)} = i_p$, and equally $\mathbf{A}^*$ and $\mathbf{S}^*$, for all events of the form $a_p^* = k_p$ and $s_{p,q}^* = j_{p,q}$ respectively.

For instance we know that because of the frame-bound priority rule, no customer in group K can be served if there are still higher priority customers unserved in group K. Formally this means that all i_j where $j < t$ must equal 0 and all i_j where $j > t$ must equal k_j. If either is not fulfilled, the above probability will equal zero. Naturally we have the condition that $1 \le i_t \le k_t$ which means there can't be more type-t customers in group K at the beginning of slot I than there were when initiating group K's service. Since $T = t$, i_t must at least be 1. From $\mathbf{R}$ and $\mathbf{A}^*$ we know that the slot I must lie within the service time of the $(k_t - i_t + 1)$'th type-t customer in group K. Because all customers of higher priority types will have been served by then, we must have that $\alpha \le n < \omega$, where

$$\alpha = \sum_{p=1}^{t-1} \sum_{q=1}^{k_p} j_{p,q} + \sum_{q=1}^{k_t - i_t} j_{t,q} \,, \qquad \omega = \sum_{p=1}^{t-1} \sum_{q=1}^{k_p} j_{p,q} + \sum_{q=1}^{k_t - i_t + 1} j_{t,q} \,. \tag{37}$$

Under the above conditions we find

$$\Pr[f^{(l)} = n, \mathbf{R} | \mathbf{A}^*, \mathbf{S}^*, T = t] = \left(\sum_{q=1}^{k_t} j_{t,q} \right)^{-1} , \tag{38}$$

$$\Pr[\mathbf{A}^*, \mathbf{S}^* | T = t] = \frac{\sum_{q=1}^{k_t} j_{t,q}}{\rho_t l / (1 - A(\mathbf{0})^l)} \Pr[b_1^{(l)} = k_1, \ldots, b_N^{(l)} = k_N] \prod_{p,q} s_p(j_{p,q}) \,, \tag{39}$$

where we used $s_p(j_{p,q}) \triangleq \Pr[s_{p,q} = j_{p,q}]$ and the product goes over the service times of all customers in group K. Equation (38) is the result of the random nature of slot I. Notice that we used a_j^* for group K's type-j population, instead of $b_j^{(l)}$, which we use for a random group, and equally $s_{j,i}^*$ instead of $s_{j,i}$. Indeed group K is not a random group, but a group whose service time contains a random slot. In this sense larger groups will be more privileged to be selected by choosing a random slot. Similarly, when we condition on type-t customers' service times, we have that groups containing type-t customers with large combined service times will be more likely to be selected as group K. As such the probability $\Pr[\mathbf{A}^*, \mathbf{S}^* | T = t]$ is proportional to $\sum_{q=1}^{k_t} j_{t,q}$. This phenomenon is also known as the renewal paradox (see f.i. [8]).

This explains equation (39). To calculate $\Pr[f^{(l)} = n, \mathbf{R}, \mathbf{A}^*, \mathbf{S}^*, T = t]$ we still need $\Pr[T = t]$ which is we reasoned to be ρ_t/ρ in the section about preliminary results above.

It remains to calculate a closed-form formula for $H_t^{(l)}(\mathbf{x}, z)$, by substituting the probabilities we found above.

$$
\begin{aligned}
H_t^{(l)}(\mathbf{x}, z) = & \sum_{k_1,\ldots,k_N=0}^{\infty} \frac{\Pr[b_1^{(l)} = k_1, \ldots, b_N^{(l)} = k_N]}{\rho l/(1 - A(\mathbf{0})^l)} \left(\prod_{p=t+1}^{N} x_p^{k_p} \right) \sum_{i_t=1}^{k_t} x_t^{i_t} \\
& \times \sum_{j_{1,1}=1}^{\infty} s_1(j_{1,1}) \cdots \sum_{j_{N,k_N}=1}^{\infty} s_N(j_{N,k_N}) \sum_{n=\alpha}^{\omega-1} z^n \\
= & \sum_{k_1,\ldots,k_N=0}^{\infty} \frac{\Pr[b_1^{(l)} = k_1, \ldots, b_N^{(l)} = k_N]}{\rho l/(1 - A(\mathbf{0})^l)} \left(\prod_{p=t+1}^{N} x_p^{k_p} \right) \sum_{i_t=1}^{k_t} x_t^{i_t} S_t(z)^{k_t - i_t} \\
& \times \left(\prod_{p=1}^{t-1} S_p(z)^{k_p} \right) \sum_{j_{t,k_t-i_t+1}=1}^{\infty} \sum_{n=0}^{j_{t,k_t-i_t+1}-1} s_t(j_{t,k_t-i_t+1}) z^n \\
= & \sum_{k_1,\ldots,k_N=0}^{\infty} \frac{\Pr[b_1^{(l)} = k_1, \ldots, b_N^{(l)} = k_N]}{\rho l/(1 - A(\mathbf{0})^l)} \left(\prod_{p=t+1}^{N} x_p^{k_p} \right) \left(\prod_{p=1}^{t-1} S_p(z)^{k_p} \right) \\
& \times x_t \, \frac{x_t^{k_t} - S_t(z)^{k_t}}{x_t - S_t(z)} \, \frac{S_t(z) - 1}{z - 1} \, . \qquad (40)
\end{aligned}
$$

Introducing $C_t^{(l)}(\mathbf{x}, z)$ we can rewrite the above in the short formula,

$$
\begin{aligned}
C_t^{(l)}(\mathbf{x}, z) \triangleq\; & B^{(l)}(S_1(z), \ldots, S_{t-1}(z), x_t, x_{t+1}, \ldots, x_N) \\
& - B^{(l)}(S_1(z), \ldots, S_{t-1}(z), S_t(z), x_{t+1}, \ldots, x_N)
\end{aligned}
$$

$$
H_t^{(l)}(\mathbf{x}, z) = x_t \, \frac{1 - A(\mathbf{0})^l}{\rho l} \, \frac{S_t(z) - 1}{z - 1} \, \frac{C_t^{(l)}(\mathbf{x}, z)}{x_t - S_t(z)} \, . \qquad (41)
$$

Notice the presence of x_t, which suggest that $r_t^{(l)} > 0$, which we already pointed out. On top of that, $H_t^{(l)}(\mathbf{x}, z)$ is no function of x_j for $j < t$, meaning that all type-j customers of group K have already been served while serving the tagged type-t customer (which is exactly the purpose of frame-bound priority). Equation (32) translates into the z-domain resulting in

$$V_a(\mathbf{z}) \triangleq E\left[\prod_{j=1}^{N} z_j^{v_j} \;\middle|\; v_T > 0,\ \hat{u} > 1\right]$$
$$= \sum_{t=1}^{N} H_t^{(M)}(\mathbf{z}, A(\mathbf{z}))\ \frac{\big(\mathbf{U}(B^{(M)}(\mathbf{z})) - \boldsymbol{\tau} B^{(M)}(\mathbf{z})\big)}{(1-\boldsymbol{\tau}\mathbf{1})B^{(M)}(\mathbf{z})^2}\mathbf{Y}^T(\mathbf{z})\ , \tag{42}$$

where we used (31) to be able to use to the vector pgf found in the previous section. $\mathbf{Y}^T(\mathbf{z})$ is a column vector with l'th element equal to $A(\mathbf{z})^l$, representing the customers that entered the system during the part of the frame preceding the k'th group departure (that are not included in $\hat{u}$). The random variable v_T is defined as the sum $\sum_{j=1}^{N} v_j$. Likewise for the two remaining cases we find

$$V_b^{(l)}(\mathbf{z}) \triangleq E\left[\prod_{j=1}^{N} z_j^{v_j} \;\middle|\; v_T > 0,\ \hat{u} = 1, \hat{m} = l, a_g^{(l)} = 1\right] = \sum_{t=1}^{N} H_t^{(l)}(\mathbf{z}, A(\mathbf{z})) \tag{43}$$
$$V_c^{(l)}(\mathbf{z}) \triangleq E\left[\prod_{j=1}^{N} z_j^{v_j} \;\middle|\; v_T > 0,\ \hat{u} = 1, \hat{m} = l, a_g^{(l)} = 0\right] = \sum_{t=1}^{N} H_t^{(1)}(\mathbf{z}, A(\mathbf{z}))\ . \tag{44}$$

Notice that $V_c^{(l)}(\mathbf{z}) = V_b^{(1)}(\mathbf{z})$. Furthermore, applying Little's result on the server we obtain $\Pr[v_T = 0] = 1 - \rho$ and so an exact expression for $V(\mathbf{z})$ has the form

$$\begin{aligned} V(\mathbf{z}) = (1-\rho)\ +\ &\rho \Pr[\hat{u} > 1 | v_T > 0] V_a(\mathbf{z}) \\ &+ \rho \sum_{l=1}^{M} \Big(\Pr[\hat{u} = 1, \hat{m} = l, a_g^{(l)} = 1 | v_T > 0] V_b^{(l)}(\mathbf{z}) \\ &\quad + \Pr[\hat{u} = 1, \hat{m} = l, a_g^{(l)} = 0 | v_T > 0] V_b^{(1)}(\mathbf{z}) \Big)\ . \end{aligned} \tag{45}$$

The unknown probabilities can be calculated as follows. If we calculate the average service time of a group - which we will denote $E[s_g]$, then we can see the above unknown probabilities as contributions to this average. This yields

$$\begin{aligned} \Pr[\hat{u} > 1 | v_T > 0] &= \frac{1 - \boldsymbol{\tau}\mathbf{1}}{E[s_g]}\ \frac{\rho M}{1 - A(\mathbf{0})^M} \\ \Pr[\hat{u} = 1, \hat{m} = l, a_g^{(l)} = 1 | v_T > 0] &= \frac{\tau_l \rho l}{E[s_g]} \\ \Pr[\hat{u} = 1, \hat{m} = l, a_g^{(l)} = 0 | v_T > 0] &= \frac{\tau_l}{E[s_g]} \frac{\rho A(\mathbf{0})^l}{1 - A(\mathbf{0})} \end{aligned} \tag{46}$$

It is easy to calculate $E[s_g]$ from the equations above since $V(\mathbf{z})$ must be normalized. This yields

$$E[s_g] = \frac{(1-\boldsymbol{\tau}\mathbf{1})\rho M}{1-A(\mathbf{0})^M} + \rho \sum_{l=1}^{M} \tau_l \left(l + \frac{A(\mathbf{0})^l}{1-A(\mathbf{0})} \right). \tag{47}$$

Now we combine the above results in (41)-(46), leading to

$$\begin{aligned} V(\mathbf{z}) = 1 - \rho \; + \; & \frac{\rho}{E[s_g](A(\mathbf{z})-1)} \sum_{t=1}^{N} \frac{z_t(S_t(A(\mathbf{z}))-1)}{z_t - S_t(A(\mathbf{z}))} \\ & \times \Bigg[\, C_t^{(M)}(\mathbf{z}, A(\mathbf{z})) \left(\frac{\mathbf{U}(B^{(M)}(\mathbf{z})) - \boldsymbol{\tau} B^{(M)}(\mathbf{z})}{B^{(M)}(\mathbf{z})^2} \right) \mathbf{Y}^T(\mathbf{z}) \\ & + \sum_{l=1}^{M} \tau_l \left[(1-A(\mathbf{0})^l) C_t^{(l)}(\mathbf{z}, A(\mathbf{z})) + A(\mathbf{0})^l C_t^{(1)}(\mathbf{z}, A(\mathbf{z})) \right] \Bigg]. \end{aligned} \tag{48}$$

For $M = 1$ and $N = 2$ this results simplifies to the result obtained in De Clercq[3], and for $M = 1$ and $N = 1$ to those found in Bruneel[2].

3.4. **Delay Distributions.** In this section we search for the delay distribution of a random type-j customer - we'll call this customer, 'customer c' - and more particularly for the probability generating functions of this drv. The delay of a random type-j customer has three parts contributing to it. First, customer c has to wait for all customers queued before it. Some of them might have entered the system during the same frame as customer c - hereafter called the entrance frame. Others will have entered during previous frames. Of the customers that entered the system during the entrance frame, the customers of higher priority classes will be queued before c as well as a fraction of those with equal priority. Lastly, after these customers are served, customer c is delayed by it's own service time s_j. Of these three components only the first two are correlated to eachother, and thus we can write:

$$D_j(z) = S_j(z) E[z^{\text{waiting time}}] \tag{49}$$

To calculate the waiting time distribution, we introduce some renewal points. Let w_k be the work in the system at the beginning of the first slot of the k'th frame. The work in the system at a specified epoch is the number of slots it takes the server to empty the content of the system (queue and server) if no new customers enter the system from the specified epoch onward. w_k forms a one-dimensional Markov chain. Under the equilibrium condition mentioned in the opening section of this paper, if we let $k \to \infty$, then w_k has a steady-state distribution and we denote its pgf by $W(z) \triangleq E[z^w]$ - in which w is the work in the system at the beginning of a random frame. We obtain this pgf by transforming the system equations for this Markov chain to the z-domain.

When $w_k = 0$, the frame that follows is of length 1 since we are in an idle period. Whenever $w_k = i$ with $0 < i < M$ however, we know the coming frame is about to get reset, and so the following frame is of length i. In all other cases, the frame that follows is of size M (full-sized). So for w_k high enough ($\geq M$), the work entering the system during the next frame is entirely uncorrelated to w_k. This yields

$$\begin{aligned} w_{k+1} &= \textstyle\sum_{j=1}^{N} \sum_{i=1}^{a_{j,1}} s_{j,i} \ , \ \text{ if } w_k = 0, \\ w_{k+1} &= \textstyle\sum_{n=1}^{l} \sum_{j=1}^{N} \sum_{i=1}^{a_{j,n}} s_{j,i,n} \ , \ \text{ if } 0 < w_k < M, \\ w_{k+1} &= \textstyle\sum_{n=1}^{M} \sum_{j=1}^{N} \sum_{i=1}^{a_{j,n}} s_{j,i,n} + w_k - M \ , \ \text{ if } w_k \geq M, \end{aligned} \tag{50}$$

where s_j's extra indices stress the fact that different customers have independent service times. When we denote the steady state probability $\Pr[w = n]$ by ω_n, transforming the system equations in the z-domain, results in

$$W(z) = \omega_0 A(\mathbf{S}(z)) + \sum_{n=1}^{M-1} \omega_n A(\mathbf{S}(z))^n + \left(W(z) - \sum_{n=0}^{M-1} \omega_n z^n \right) \frac{A(\mathbf{S}(z))}{z^M} \tag{51}$$

$$W(z) = \frac{\omega_0 z^M \left(A(\mathbf{S}(z)) - 1\right) + \sum_{n=0}^{M-1} \omega_n \left(A(\mathbf{S}(z))^n z^M - A(\mathbf{S}(z))^M z^n \right)}{z^M - A(\mathbf{S}(z))^M}. \tag{52}$$

This pgf has M unknowns namely ω_n, $n \in \{0, \ldots, M-1\}$. The normalization condition gives us one equation in these M unknowns. Applying Rouché's theorem we can see that $z^M - A(\mathbf{S}(z))^M$ has M zeros inside the closed unit disk (see f.i. Klimenok[9]) denoted by ζ_k, $k \in \{1, \ldots, M\}$. Since pgf's are bounded in the closed unit disk, each zero of the numerator must be a zero of equal multiplicity - multiplicity 1 - in the denominator as well. Solving

$$\zeta_k = A(\mathbf{S}(\zeta_k))\epsilon_M^k \ , \ \ k \in \{1, \ldots, M\}, \tag{53}$$

for the M ζ_k's, will therefore lead to a set of M linear equations in ω_0 through ω_{M-1}, namely those found by substituting z by ζ_k in the numerator in (52) and equating this numerator to zero

$$\omega_0(\zeta_k \epsilon_M^{-k} - 1) = \sum_{n=0}^{M-1} \omega_n \zeta_k^n (1 - \epsilon_M^{-kn}) \ , \ \ \forall k \in \{1, \ldots, M\}. \tag{54}$$

The M'th zero of the denominator is of course $\zeta_M = 1$, which yields a trivial equation that is replaced by the normalization condition. Gail[5] proves that if w has a limiting distribution, this system of linear equations together with the normalization condition has a unique solution.

The attentive reader, may have noticed that we did in fact not redefine ζ_k, since these quantities are also the eigenvalues of $D(G)$ (see (24)) that were used in the calculation of the boundary vector $\boldsymbol{\tau}$, and the above system of equations (again $M-1$ non-trivial ones) bares a great resemblance to (30). In fact, knowledge of one, readily translates into knowledge of the other, as will be shown next.

When at the beginning of a frame the work in the system is $n \in \{1, \ldots, M\}$, that frame will be reset n slots later in phase n because of a group departure (groups can only enter the system at the end of a frame). Naturally a frame can only be reset when one slot prior to it, the system contained only one group. Hence for every frame at the start of which the work in the system is n, there is a group departure which resets a frame, causing the previous to finish in phase n - which means $\omega_n \sim \tau_n, \ \forall n \in \{1, \ldots, M\}$. However there are frames during which no group enters the system and so there are more observation epochs for w than for (u, m). The proportionality constant is therefore given by the probability that a group enters the system during a random frame ($= \eta$).

$$\eta = \omega_0(1 - A(\mathbf{0})) + \sum_{i=1}^{M-1} \omega_i (1 - A(\mathbf{0})^i) + M(1 - \sum_{i=0}^{M-1} \omega_i)(1 - A(\mathbf{0})^M) \ , \tag{55}$$

$$\frac{\omega_n}{\tau_n} = \eta \ , \ \ \forall n \in \{1, \ldots, M\} \ . \tag{56}$$

in which it is important to note that η is independent of n. Of the above equations, only $M-1$ are useful - ω_0 is not mentioned. We can however substitute the τ_n in the system of equations (54). Through transitivity with (30) we arrive at

$$\omega_0 = \eta \sum_{l=1}^{M} \tau_l \frac{A(\mathbf{0})^l}{1 - A(\mathbf{0})} \ . \tag{57}$$

An alternative way of obtaining this last equation could equally be the observation that the work in the system at the start of a random frame can only be 0 when at the start of the previous frame the work in the system was $n \leq M$ and no customers entered the system since.

As mentioned earlier, customer c's delay is caused in part by customers that entered the system before the entrance frame that are not yet served upon customer c's arrival - hereafter referred to as the senior customers. An upper bound for this part of the delay is given by the work in the system at the beginning of the entrance frame, which we will denote by $\hat{w}$. The distribution of $\hat{w}$ is not the same as w's, since for $\hat{w}$ a random type-j customer was chosen - and by the BASTA-property, a random slot (see f.i. Halfin[7]) - where for w a random frame must be selected. The probability that a random type-j customers enters the system during a frame is proportional to the length of that frame, and so we find that the relation between w and $\hat{w}$ is as follows.

$$\begin{aligned} \Pr[\hat{w} = 0] &= \omega_0 \phi^{-1} \\ \Pr[\hat{w} = n] &= n\omega_n \phi^{-1}, \;\; 0 < n < M \\ \Pr[\hat{w} \geq M] &= M(1 - \sum_{n=0}^{M-1} \omega_n)\phi^{-1}, \end{aligned} \tag{58}$$

in which ϕ is the average frame size. The above probabilities must be normalized. Combined with the normalization condition on w, we can find that

$$\phi\,(1-\rho) = \omega_0, \tag{59}$$

which is in agreement with the observation that the probability that a type-j customers enters the system during an idle period ($\hat{w} = 0$) must be equal to $1 - \rho$.

The effective delay caused by senior customers to customer c, is strictly less than $\hat{w}$ because customer c's waiting time starts at the beginning of the slot following its arrival slot. Customer c's arrival slot can be any of the slots in the entrance frame with equal probability. Let $\hat{w}^-$ be the effective delay caused by senior customers - the minus indicating that it will be less than $\hat{w}$. Because knowledge of $\hat{w}$ gives us the length of the entrance frame, we condition $\hat{w}^-$ on $\hat{w}$ obtaining

$$\Pr[\hat{w}^- = i | \hat{w} = n] = \begin{cases} 1 & \text{if } n = 0 \text{ and } i = 0, \\ n^{-1} & \text{if } 0 < n < M \text{ and } i < n, \\ M^{-1} & \text{if } M \leq n \text{ and } n - M \leq i < n. \end{cases} \tag{60}$$

Lastly, the delay contributed by customers entering during the entrance frame is made up of the service times of all customers of higher priority, and some customers of the same priority class as customer c. It can be calculated (see f.i. [2]) that the joint pgf of the number of customers of each type entering the system during the entrance frame and queued before customer c, given that the entrance frame was n slots long, is given by

$$\tilde{A}_j^{(n)}(\mathbf{z}) \triangleq \frac{A(z_1,\ldots,z_{j-1},z_j,1,\ldots,1)^n - A(z_1,\ldots,z_{j-1},1,1,\ldots,1)^n}{n\lambda_j(z_j-1)} \tag{61}$$

And thus the pgf of the amount of work these customers stand for equals $\tilde{A}_j^{(n)}(\mathbf{S}(z))$. All these units of work contribute to the delay of customer c, since none of them can get served until after the entrance frame finishes, as agreed in section 2, i.e. after customer c's arrival. Bringing all of the above together, we find that

$$\begin{aligned} E[z^{d_j}|\hat{w}=0] &= S_j(z)\tilde{A}_j^{(1)}(\mathbf{S}(z)), \\ E[z^{d_j}|\hat{w}=n, 0<n<M] &= S_j(z)\tilde{A}_j^{(n)}(\mathbf{S}(z))\frac{z^n-1}{n(z-1)}, \\ E[z^{d_j}|\hat{w}\geq M] &= S_j(z)\tilde{A}_j^{(M)}(\mathbf{S}(z))\frac{z^M-1}{M(z-1)}\,\frac{W(z)-\sum_{i=0}^{M-1}\omega_i z^i}{z^M(1-\sum_{i=0}^{M-1}\omega_i)}. \end{aligned} \tag{62}$$

For this last equation we used $\Pr[w=n|w\geq M] = \Pr[\hat{w}=n|\hat{w}\geq M]$. The pgf with $z-1$ in the denominator is that of $\hat{w}^-$. Using the probabilities found in (58) and the pgf found for $W(z)$ in (52) we can derive the pgf of the delay of a random type-j customer $D_j(z)$:

$$\begin{aligned} D_j(z) &= \phi^{-1}S_j(z)\omega_0\left(\tilde{A}_j^{(1)}(\mathbf{S}(z)) + \tilde{A}_j^{(M)}(\mathbf{S}(z))\frac{z^M-1}{z-1}\,\frac{A(\mathbf{S}(z))-1}{z^M-A(\mathbf{S}(z))^M}\right) \\ &+ \phi^{-1}S_j(z)\sum_{n=1}^{M-1}\omega_n\left(\tilde{A}_j^{(n)}(\mathbf{S}(z))\frac{z^n-1}{z-1} + \tilde{A}_j^{(M)}(\mathbf{S}(z))\frac{z^M-1}{z-1}\,\frac{A(\mathbf{S}(z))^n-z^n}{z^M-A(\mathbf{S}(z))^M}\right). \end{aligned} \tag{63}$$

4. **Numerical Results.** In an attempt to show the effect of the proposed FBP scheduling discipline, we introduce a composite Poisson arrival process with two types of customers. The two-dimensional ($N=2$) pgf $A(z_1,z_2)$ is chosen to be

$$A(z_1,z_2) = e^{\lambda(p_1z_1+p_2z_2-1)}, \tag{64}$$

in which λ is the workload and arrival rate (we will chose service times deterministically equal to 1 slot), and each Poisson event generates a type-j customer with probability p_j, $j\in\{1,2\}$. Since, for example, in data networks high-priority customers (control messages) are more scarce than low-priority customers (raw data), we choose $p_1=0.1$ and $p_2=0.9$. Notice that with the above pgf both arrival streams are independent of one another.

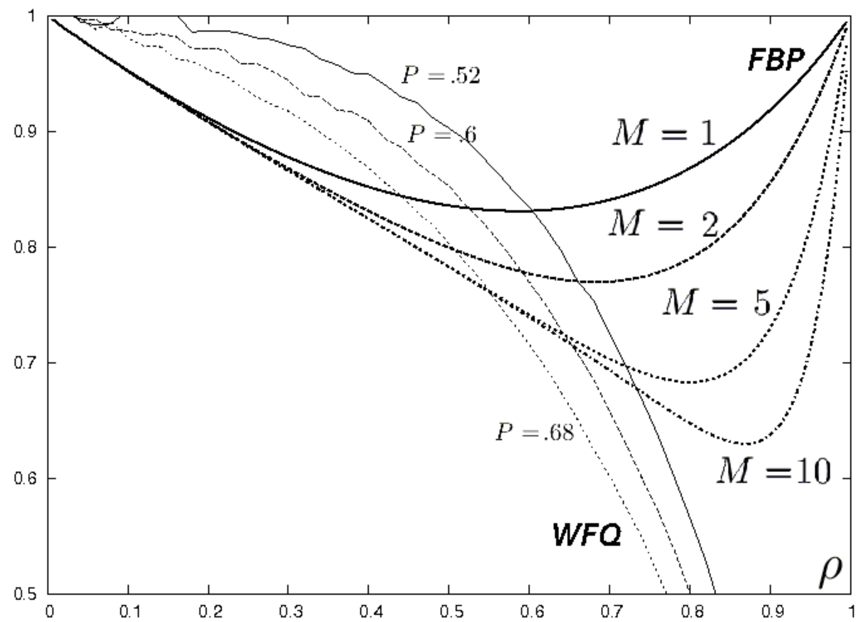

FIGURE 5. delay differentiation between type-1 and type-2 customers ($E[d_1]/E[d_2]$) in a system under frame-bound priority - for various (maximum) frame-sizes - and in a system under weighted fair queueing - for various weights of the high priority traffic ($= P$) - plotted against ρ. Data points for WFQ were obtained via simulation.

We plot the average customer delay, which for real-time applications is one of the main performance measures. These can be obtained as the first derivatives of either $D_j(z)$ or $V(\mathbf{z})$, through Little's result. A good measure for delay differentiation would be $E[d_1]/E[d_2]$, since it tells us what effect a higher priority can cause on average. In Fig.5 this measure was plotted against the workload ρ for different (maximum) frame lengths ($M = 1, 2, 5$ and 10). The same arrival stream was fed into a similar system, where the service discipline was replaced by weighted fair queueing (WFQ) in which the weight of the high-priority packets was varied ($P = .52, .6$ and $.68$) [1]. The results of these simulations are also plotted in the same figure, for comparison purposes.

We observe that for low loads tuning M in FBP has little effect on the delay differentiation - i.e. what percentage type-1 customers are delayed compared to type-2 customers. This is because, depending on the load for high enough M, the vast majority of frames will reset before completion, rendering the exact value of M somewhat irrelevant. One by one the curves pertaining to $M = 1, 2, 5$ and 10, show different values for increasing ρ, signaling that for those and higher values of the load, frames of maximum size are formed and queued. Eventually the curve for each M reaches a maximal delay differentiation (i.e., a minimal value of $E[d_1]/E[d_2]$) for a different load, and for higher M this maximum corresponds to higher loads. For

[1] In our simulations, we adopted a discretized version of WFQ

$M = 10$ f.i. this optimum almost reaches 60%, but it is obviously very load-sensitive. In the limit ($\rho \to 1$) every curve turns back to 1. This is because for every finite M, if the load is chosen high enough, many groups will typically accumulate in the system, and the customer delay of any type is primarily determined by the service times of the groups that are queued before a tagged customer's group, rather than by the ordering of customers within this customer's group. Therefore, different types of customer experience comparable delays, provided that ρ lies close enough to 1 and M is finite.

An essential difference between WFQ and FBP is the limiting cases ($P = 1$ and $M \to \infty$). WFQ scheduling evolves to absolute priority when the weight of the high priority class customers goes to 1, whereas for infinitely large frames, in FBP a kind of 'gated' priority is obtained in which frames last until they are at last reset. In the FBP case, the delay differentiation remains limited, even for infinite M and $\rho \to 1$. Also, for WFQ, scenarios may occur in which the average delay of high priority customers exceeds that of lower priority customers, i.e. when there's a lot of high priority traffic ($E[d_1]/E[d_2] > 1$). WFQ was engineered this way so that the individual streams can only punish themselves for being greedy. As a consequence, when the high priority customers pertain to different sources, they can get punished unrightfully. Notice that contrary to WFQ, the proposed scheduling discipline FBP does not give low priority traffic a lower average delay unless for very specific settings for M and the arrival process such as correlation between the number of arrivals of different types of customers during the same slot and the workload. Another key difference between WFQ and FBP, lies in the light traffic behaviour of the two, in which case WFQ's tunability is much larger. Unless traffic is very bursty (such that $b_j^{(1)} \approx b_j^{(M)}$ even for high M), this is not an issue.

Typical figures show that the power of FBP lies in the property that the delay differentiation can be controlled (of course within certain bounds), if the load is medium to high (i.e. .5 - .95), as can also be observed in Fig.6. This figure is the result of applying a bursty arrival process with joint pgf $A(z_1, z_2) = \frac{9}{10} + \frac{1}{10}e^{10\lambda(p_1 z_1 + p_2 z_2 - 1)}$. In this scenario, very few slots experience customer arrivals, but the few that do see (on average) 10 times as much arrivals compared to the previous figure. As already hinted, increasing M not too much has very little effect; however as can be seen for larger M, we can still tune the delay differentiation quite well for moderate to high values of the load. On the other hand, when M approaches infinity, the delay differentiation becomes almost independent of the load, unless the load is close to 0 or 1.

Using these observations together with an expression for the first derivative of $D_j(z)$ (see (63)), one can tune the frame length to obtain the desired QoS-level, for instance based on the results projected in a figure such as Fig.6. The FBP scheduling discipline is then easily implemented using a timer and a reorder buffer.

5. **Conclusions.** In this paper we introduced and analyzed a new scheduling mechanism, called frame-bound priority, in order to give an answer to the demand for delay differentiation between different traffic flows (requiring different QoS). The proposed scheduling mechanism, which partitions time into consecutive frames and

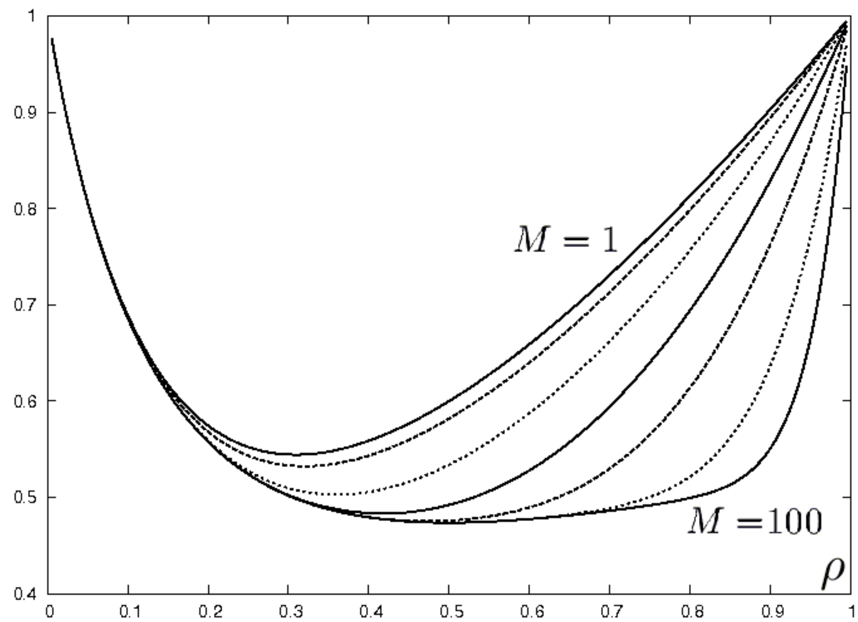

FIGURE 6. $E[d_1]/E[d_2]$ in a system under frame-bound priority for various (maximum) frame-sizes, with a bursty arrival process. ($M = 1, 2, 5, 10, 20, 50$ and 100)

allows high priority customers to overtake lower priority customers that entered the system during the same frame, was analyzed using a combination of matrix analytic methods and (joint) generating functions on a discrete-time queueing model. Our results allow to calculate the joint probability generating function of the number of customers of all classes separately, given any i.i.d. arrival process. In addition, we have also examined in detail the effect frame-bound priority has on delays of the individual customer classes, both by analysis and some basic numerical examples. Thanks to an efficient method for calculating the boundary vector of the underlying Markov chain, data points can be generated both accurately and fast. The numerical examples show that tuning of the frame length provides a flexible way to achieve delay differentiation between the customer classes, all while solving the starvation problem, i.e. low priority traffic cannot be delayed indefinitely by design.

REFERENCES

[1] D.A. Bini, G. Latouche and B. Meini "Numerical Methods for Structured Markov Chains", Oxford University Press, 2005, ISBN: 0-19-852768-3.

[2] H. Bruneel *Performance of Discrete-Time Queueing Systems*, Computers Operations Research, **20** (1993), 303–320

[3] S. De Clercq, B. Steyaert and H. Bruneel *Analysis of a Multi-Class Discrete-time Queueing System under the Slot-Bound Priority rule*, Proceedings of the 6th St. Petersburg Workshop on Simulation, (2009), 827–832.

[4] S. De Vuyst, S. Wittevrongel and H. Bruneel *A queueing discipline with place reservation*, Proceedings of the COST 279 Eleventh Management Committee Meeting (Ghent, 23-24 September 2004), COST279TD(04)36.

[5] H.R. Gail, S.L. Hantler and B.A. Taylor *Spectral Analysis of M/G/1 and G/M/1 Type Markov Chains*, Advances in Applied Probability, **28** (1996), 114–165.
[6] R.G. Gallager "Discrete Stochastic Processes", Kluwer Academic publishers, 1996.
[7] S. Halfin *Batch Delays Versus Customer Delays*, The Bell System Technical Journal, **62** (1983), 2011–2015.
[8] L. Kleinrock "Queueing Systems, Volume I: Theory", Wiley, New York, 1975.
[9] V. Klimenok *On the modification of Rouche's Theorem for Queueing Theory Problems*, Queueing Systems, **38** (2001), 431–434
[10] K.Y. Liu, D.W. Petr, V.S. Frost, H.B. Zhu, C. Braun and W.L. Edwards *Design and Analysis of a Bandwidth Management Framework for ATM-Based Broadband ISDN*, IEEE Communications Magazine, **35** (1997), 138–145.
[11] M.F. Neuts "Structured Stochastic Matrices of M/G/1 Type and Their Applications", vol.5 of Probability: Pure and Applied. Marcel Dekker Inc., New York, 1989, ISBN: 0-8247-8283-6.
[12] I. Stavrakakis *Delay Bounds on a Queueing System with Consistent Priorities*, IEEE Transactions on Communications, **42 No.2-4** Part 1 (1994), 615–624
[13] H. Takada and K. Kobayashi *Loss Systems with Multi-Thresholds on Network Calculus*, Proc. of Queueing Symposium (2007), 241–250.
[14] J. Walraevens, B. Steyaert and H. Bruneel *Performance Analysis of The System Contents in a Discrete-Time Non-Preemptive Priority Queue with General Service Times*, Belgian Journal of Operations Research, Statistics and Computer Science (JORBEL), **40** (2001), 91–103
[15] H.S. Wilf "Generatingfunctionology", Department of Mathematics, University of Pennsylvania, Academic Press Inc., 1994, ISBN: 0-12-751956-4.

Received September 2010; 1st revision January 2011; final revision May 2011.

E-mail address: Sofian.DeClercq@telin.ugent.be
E-mail address: Koen.DeTurck@telin.ugent.be
E-mail address: Bart.Steyaert@telin.ugent.be
E-mail address: Herwig.Bruneel@telin.ugent.be