Performance of the GFR-Service with Constant Available Bandwidth

Norbert Vicari and Robert Schedel

Institute of Computer Science, University of Würzburg,
Am Hubland, D - 97074 Würzburg, Germany
Tel.: +49 931 888 5505, Fax.: +49 931 888 4601
e-mail: [vicari|schedel]@informatik.uni-wuerzburg.de

Abstract

Recently the ATM Forum defined the Guaranteed Frame Rate service category to provide a minimum service guarantee to classical best-effort services. The eligibility of frames for this service category is determined with the Frame Based Generic Cell Rate Algorithm and the transmission is guaranteed by a queuing-discipline based on two thresholds. In this paper we present a discrete-time analysis of the GFR-service under the assumption of constant available bandwidth. The presented method can be applied to dimension the thresholds of the queuing discipline used to enforce the guaranteed service.

1 Introduction

The ATM-Forum recently introduced the *Guaranteed Frame Rate* (GFR) service category [3]. This new service category is motivated by several intentions. Today most applications are not equipped to select the proper traffic parameters required to establish ATM connections. Thus, choosing CBR or VBR service categories will fail either by causing inefficiency by overestimating required resources or not being able to give any QoS guarantees. The ABR service category is regarded to be too complex to be implemented in the majority of systems. Further, the efficiency of the ABR control loop is reduced by increasing distance of source and destination.

Transferring data traffic with the best-effort service category UBR would avoid the problem of estimating traffic descriptors, but will also give no QoS guarantees at all. Worse, the throughput seen at higher protocol layers is severely reduced. Most of todays applications utilize the *Transmission Control Protocol* (TCP) for transferring data in frame based structures. When transmitting these frames over an ATM network the data is fragmented into cells. The loss of a single cell will cause an irrecoverable damage to the whole frame and induce retransmission. To cope with these problems, the GFR service category provides the user with a *Minimum Cell Rate* (MCR) guarantee under the assumption of a given *Maximum Frame Size* (MFS) and *Maximum Burst Size* (MBS). The user is allowed to send traffic in excess of the negotiated parameters, but this traffic will only be delivered within the limits of available resources.

To offer the guarantees of the GFR service category the functions of the ATM traffic management have to distinguish frames sent in line with the negotiated parameters and excess traffic. The frames are classified applying the *Frame Based Generic Cell Rate Algorithm* (F-CGRA), which decides the eligibility of frames for the GFR service. Similar to the conventional GCRA the state of the F-GCRA is changed with every arriving cell, but only whole frames are regarded or disregarded as eligible for the GFR service.

In order to offer GFR service it is not only important to be able to distinguish eligible and non-eligible frames, but also to discard cells properly. Discarding all cells belonging to non-eligible frames would be GFR service compliant, but would offend the best-effort nature of the service. Thus, the queuing behavior of the GFR service has to be studied. A trade-off between providing guarantees for eligible frames and offering best-effort service for excess traffic has to be encountered.

The paper is organized as follows: in Section 2 an overview of the GFR-service category and its key components is given. Section 3 describes the modeling and analysis of the system. Numerical examples derived with the presented analysis method are provided in Section 4. The paper is concluded with a summary.

2 System Description

Motivated by the needs of a guaranteed minimum bandwidth for best-effort ATM connections the introduction of a MCR for the UBR service category was suggested [1]. This considerations resulted in the definition of the so called UBR+ service category [2]. The main idea of this service category was to preserve the best-effort properties of the UBR service category while adding the guarantee of a minimum bandwidth. The newer specification of this service category [3] names the service GFR, which reflects the approach of taking frames into account for the minimum guaranteed bandwidth.

In comparison to ABR, which also provides a guaranteed best-effort service, GFR is easier to implement and does not add a new flow control scheme. Thus, implementation of GFR in adapter cards and network nodes is expected to be faster and cheaper. Further, the coupling of different flow control schemes – like TCP over ABR – may lead to unpredictable and unintended complications.

The VBR.3 service category [4] also allows the user to send traffic in excess to the traffic contract, but in comparison to the GFR service the traffic is not regarded as flow of frames. Since, most currently available applications use the TCP protocol, data is organized in frames and has to be fragmented into cells for transport over an ATM net. Thus, random loss of a single cell leads to corruption of the whole frame and reduces the goodput of the transmission. Discarding whole frames that are not eligible for guaranteed transmission increases the goodput of the net.

The GFR service is intended to support non-real-time traffic and requires the data being organized in frames which can be delineated at the ATM layer. The user is provided with a MCR guarantee when transmitting frames that do not exceed the MFS in a burst that does not exceed the MBS. Frames sent in excess to this parameters will be delivered only within the limits of available resources.

The user can indicate excess traffic by marking frames, that is setting CLP = 1 in every cell of a frame. This frame will be regarded of less importance and the MCR guarantee will not apply to marked frames. If the user has requested the tagging option, the network is supposed to mark non-eligible frames and transport them without MCR guarantee.

The GFR service [5] provides a guarantee to deliver complete unmarked frames that are *conforming* and *eligible*. A frame is defined to be conforming if the CLP bit of all cells of a frame has the same value as the CLP bit of the first cell of the frame, the number of cells on the frame is less than MFS and the rate of the cells conforms to the parameter *Peak Cell Rate* (PCR), which is monitored with a conventional GCRA. The eligibility of frames is defined with the F-GCRA, which controls if the rate of the cells of a frame is less than MCR and the length of burst is less than MBS. In order to stick to the frame-based approach all cells of a frame are valued identically to the first cell of a frame.

In order to provide the above defined service guarantee the network has to discriminate against eligible and non-eligible frames when transmitting data with the GFR service. Thus we will review the queuing discipline applied for the GFR service after introducing the exact functionality of the F-GCRA.

2.1 Frame-Based Generic Cell Rate Algorithm (F-GCRA)

Figure 1 shows how the *F-GCRA(I,L)* algorithm decides if an arriving frame is eligible or not according to a given increment parameter *I* and a limit parameter *L*. The only basic difference of *F-GCRA(I,L)* to a *GCRA(I,L)* – as formerly been defined by the ATM traffic management standard [4] – is that only the first cell of a frame is checked according to the GCRA while effectively tagging the whole frame as eligible or non-eligible. Frame cells arriving later only update the F-GCRA state if the first cell was eligible but they are never checked against the limit parameter *L*. In other words, a conventional *GCRA(I,L)* is exactly equivalent to a *F-GCRA(I,L)* if all arriving frames have a constant size of one cell. Note that frames whose cell loss priority bit is set (CLP = 1) are always regarded non-eligible – analogously to the behavior of the traditional GCRA.

```
INPUT cell arrival on time time
temp = X - (time - last_eligible_time)
if (first cell in frame) then
          if  (CLP == 1) or (temp > L) then
                    frame_eligible = false
          else
                    frame_eligible = true
          endif
endif

if (frame_eligible) then
          X  = max(0, temp) + I
          last_eligible_time = time
endif

OUTPUT  cell eligible according to  F-GCRA(I,L) if (frame_eligible = true)
```

Figure 1: *F-GCRA(I,L)* implemented as virtual scheduling algorithm.

The simplified F-GCRA algorithm above assumes that the frame stream is already conforming to the *PCR* with frame sizes no bigger than *MFS*. For the treatment of non-conforming frames no binding rule exists. However, it is very common practice to regard a non-conforming frame as a strict violation of the traffic contract with the immediate consequence of complete frame rejection. Hence, for the addressed GFR performance evaluation an inclusion of further details in the algorithm is neither necessary nor sensible. Figure 2 shows an illustration of the F-GCRA state for an example scenario with three consecutively arriving frames.

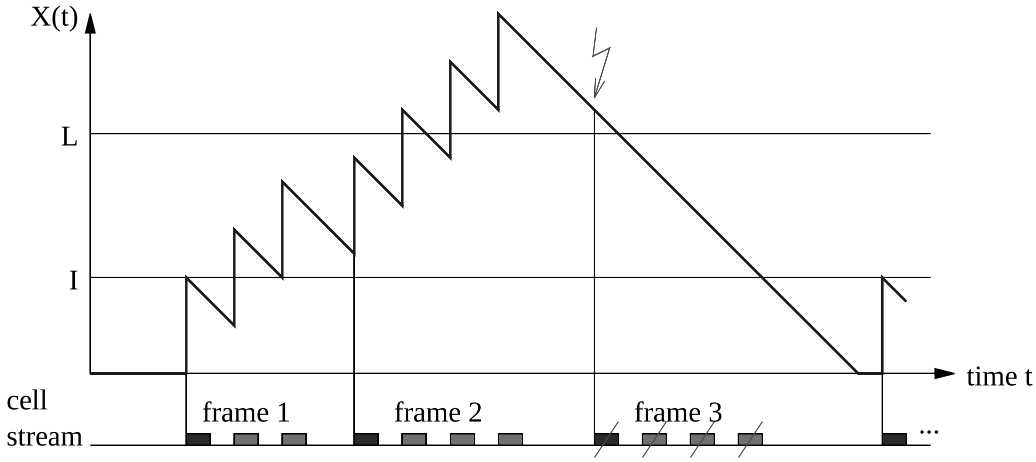

Figure 2: F-GCRA state diagram with three arriving frames.

The first two frames are declared to be eligible according to *F-GCRA(I,L)* as the F-GCRA state upon arrival of the first cell does not exceed the limit parameter *L*. For each arriving cell of these two frames the F-GCRA state is increased by the increment parameter *I*. The third frame is declared non-eligible and the F-GCRA is never increased upon arrival of any frame cell.

2.2 Queuing Discipline

After classifying the frames with help of the F-GCRA the network node has to transmit the frames eligible for service guarantee with low loss probability. If additional resources are available on the transmission link, frames sent in excess to the traffic contract should be also transferred. Naturally these frames will suffer a higher loss rate than frames with guaranteed service. Cells which could not be transferred immediately are stored in a buffer of size *Q_MAX*. When a cell of a frame – eligible or non-eligible – could not be stored in the buffer, this cell and all subsequent cells of this frame are discarded since it is assumed that the loss of a single cell of a frame leads to the retransmission of the whole frame.

To discern eligible and non-eligible frames two threshold values are introduced. The Low Buffer Occupancy (*LBO*) value indicates the limit for the acceptance of non-eligible frames. That is, if at the time instant of the arrival of the first cell of a non-eligible frame at least *LBO* cells are waiting in the buffer the whole frame is discarded. Analogously the High Buffer Occupancy *(HBO*) value defines the limit for the acceptance of eligible frames. Once the first cell of a frame is accepted the subsequent cells of this frame could be only discarded due to buffer overflow. In order to investigate the influence of different values of *LBO* and *HBO* we model the correlated system of F-GCRA and queue.

3 Model and Analysis

In the following, a model and its corresponding discrete-time analysis of the GFR service are presented. The issue of conforming/non-conforming frames will be neglected. We will also assume, that the transmission capacity for the GFR-service is constant. After a description of the frame arrival process we will evaluate the system state of the coupled model of F-GCRA and transmission line as shown in Figure 3.

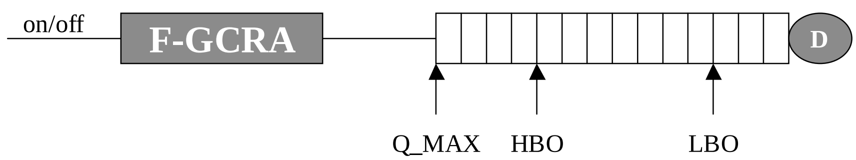

Figure 3: GFR Model

For the analysis of the model we will describe the system state by a two dimensional random variable (X_f, X_q). The first dimension represents the counter of the leaky bucket of the F-GCRA and the second dimension represents the time required to transmit the cells waiting in the buffer.

All random variables to describe the system state are measured in multiples of the duration of a cell transmission at PCR. The constant capacity of the transmission link is denoted by the RV T_S, that is, at most a cell could be transmitted every T_S time units. Consequently the capacity QL of the buffer including the transmission unit computes as:

$$QL = (Q\_MAX + 1)T_S - 1 . \tag{1}$$

Analogously the limits for the acceptance of eligible and non-eligible frames HBL and LBL – expressed in time-slots – are defined:

$$\begin{aligned} LBL &= (LBO + 1)T_S - 1, \\ HBL &= (HBO + 1)T_S - 1. \end{aligned} \tag{2}$$

An analysis of the F-GCRA for determining the parameters MCR and MBS can be found in [6].

3.1 Modeling the Arrival Process

The functionality of the GFR service is based on the organization of data in frames. Most of the currently used protocols e.g. TCP/IP transport data in frames that have to be split up in several cells for transport over an ATM network. This sort of traffic can be modeled by the class of on/off-processes. An on-state represents the transmission of a frame, while the off-state re-presents the time between the frames. In our analysis we consider an on/off source as depicted in Figure 4.

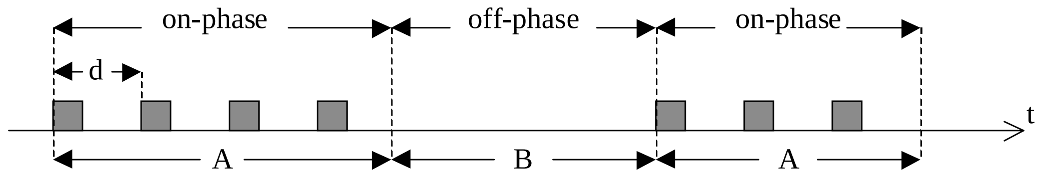

Figure 4: On/off source.

The duration of the on- and off-phases are distributed according to discrete general and independent distributions. At the beginning of a on-phase a cells arrives immediately. During an on-phase cells arrive in intervals of *d* time units, which correspond to the transmission time of a cell *1/PCR*. The end of an on-phase is not generally synchronized to a cell arrival.The duration of the on-phase is denoted by the random variable A and the duration of the off-phase by B, respectively.

3.2 System State Evolution

The random variable (X_f^{on}, X_q^{on}) describes the system state at the beginning of an on-phase. Tracing every cell arrival the system state after the arrival of the last cell of a frame is iteratively derived. Taking into account the time remaining in the on-phase the system state at the beginning of the off-phase (X_f^{off}, X_q^{off}) is given. Since no cell arrivals occur during the off-phase the system state is decreased until the beginning of the next on-phase is reached.

3.2.1 Transition: Off-Phase to On-Phase

The state of the F-GCRA is reduced during the off-phase by *B* units, since no cell arrival occurs. Before the arrival of the next frame the state of the F-GCRA computes as follows:

$$X_f^{on} = max(X_f^{off} - B, 0) . \tag{3}$$

Analogously at most B cells could be served at PCR during the off-phase, thus the state of the queue at the beginning of the on-phase is given by the following equation:

$$X_q^{on} = max(X_q^{off} - B, 0)\,. \tag{4}$$

3.2.2 Transition: On-Phase to Off-Phase

During the on-phase cell arrivals occur starting with the first time slot and continuing every *d* time slots, c.f. Figure 4. The system state is denoted recursively, and the system is observed at the end of time slot *k*.

At the beginning of the on-phase the F-GCRA takes the following state:

$$X_f^{on,0} = X_f^{on}\,. \tag{5}$$

The state of the F-GCRA at the end of an on-phase of *A* time units is computed recursively as follows:

$$X_f^{on,A} = \begin{cases} X_f^{on,A-1} - 1 + I, & \text{if } X_f^{on,0} \le L \text{ and cell arrival} \\ max(0, X_f^{on,A-1} - 1), & \text{else} \end{cases}\,. \tag{6}$$

For the derivation of the buffer state we discern the arrival of the first and the subsequent cells of a frame. The state of the buffer before the arrival of the first cell is given by:

$$X_q^{on,0} = X_q^{on}\,. \tag{7}$$

Upon the arrival of the first cell of a frame the acceptance of the frame is decided. If the state of the buffer at most *LBL* all frames are accepted. Eligible frames are accepted even if the state of the buffer is higher then *LBL* but at most *HBL*. If a cell could not be accepted all remaining cells of the frame are also discarded. We denote this by changing the value of the RV *DF* from 0 to 1.

$$X_q^{on,1} = \begin{cases} X_q^{on,0} - 1 + T_{S,} & \text{if } X_f^{on,0} \le L \text{ and } (X_q^{on,0} \le HBL - T_S) \\ X_q^{on,0} - 1 + T_{S,} & \text{if } X_f^{on,0} > L \text{ and } (X_q^{on,0} \le LBL - T_S) \\ max(0, X_q^{on,0} - 1), & \text{else} \end{cases}\,. \tag{8}$$

The remaining cells of a frame are discarded only if either the buffer is occupied or some cells of the frame have already been discarded. Thus the state of the buffer is recursively computed as follows:

$$X_q^{on, A} = \begin{cases} X_q^{on, A-1} - 1 + T_{S,} & \text{if } X_q^{on, A-1} \leq QL - T_{S,} \text{ and } DF = 0 \text{ and cell arrival} \\ max(0, X_q^{on, A-1} - 1), & \text{if } DF = 1 \text{ and cell arrival} \\ max(0, X_q^{on, A-1} - 1), & \text{if not cell arrival} \\ max(0, X_q^{on, A-1} - 1), & \text{else} \end{cases} \quad . \tag{9}$$

The 'else' branch of Equation (8) and (9) indicates the first discarding of a cell and thus initiates a change of the value of *DF* from 0 to 1.

3.3 Operators of the discrete-time analysis

Before we derive the probability mass functions describing the system state, we introduce some operators to simplify the description.

To represent the probability mass function corresponding to $max(m, X)$, that is the maximum of a constant value *m* and a random variable *X* the projection operator π is defined. Since, we use two-dimensional random variables in this investigation the projection is defined separately for both dimensions:

projection of the first dimenson: projection of the second dimenson:

$$\pi_{1, m}[x(i, j)] = \begin{cases} 0 & : i < m \\ \sum_{k = -\infty}^{m} x(k, j) & : i = m \\ x(i, j) & : i > m \end{cases}, \qquad \pi_{2, m}[x(i, j)] = \begin{cases} 0 & : j < m \\ \sum_{k = -\infty}^{m} x(i, k) & : j = m \\ x(i, j) & : j > m \end{cases} . \tag{10}$$

In order to split a distribution into two semidistributions the split operator σ is defined. In this investigation we apply the operator with regard to the second dimension of the distribution describing the system state:

lower part: upper part:

$$\sigma_m[x(i, j)] = \begin{cases} x(i, j) & : j \leq m \\ 0 & : j > m \end{cases}, \qquad \sigma^m[x(i, j)] = \begin{cases} 0 & : j \leq m \\ x(i, j) & : j > m \end{cases} . \tag{11}$$

3.4 Derivation of the corresponding probability mass functions

To express the systems memory regarding the loss of subsequent cells of a frame, we have to split the distributions representing the system state.

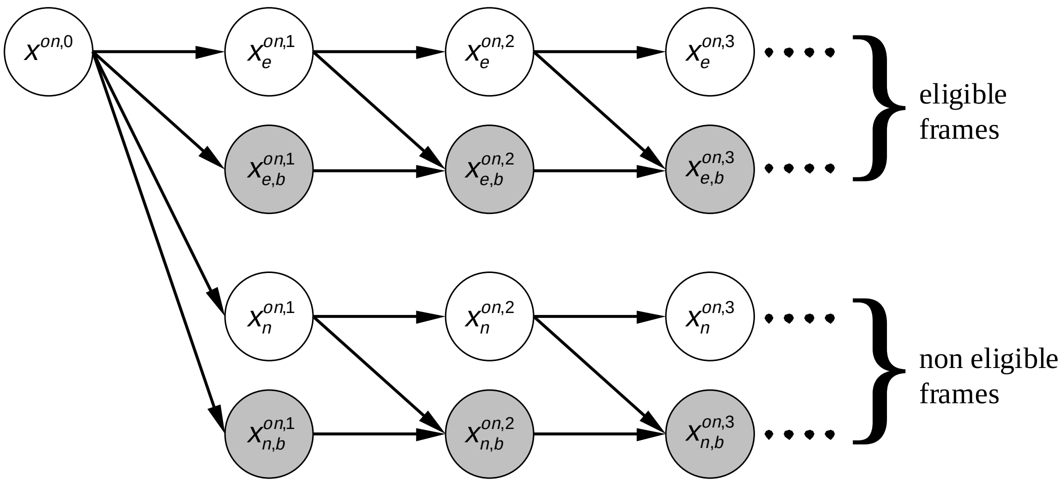

Figure 5: Transitions between the distributions describing the system state

Figure 5 depicts the transitions between the distributions of the system state. After the arrival of the first cell of a on-phase the system state is described by four semi-distributions – eligible frames, discarded eligible frames, non-eligible frames, discarded non-eligible frames. Discarding cells is represented by shifting a part of the distribution which represents accepted cells to the distributions representing discarded cells. In the following we describe exactly the evolution of the distributions.

For $k = 1$ we obtain the following relations:

$$
\begin{aligned}
x_e^{on,1}(f-1+I, q-1+T_S) &= \sigma_{HBL-T_S}[x^{on,0}(f,q)] & f &= 0, 1, \ldots, L, \\
x_{e,b}^{on,1}(f-1+I, q-1) &= \sigma^{HBL-T_S}[x^{on,0}(f,q)] & f &= 0, 1, \ldots, L, \\
x_n^{on,1}(f-1, q-1+T_S) &= \sigma_{LBL-T_S}[x^{on,0}(f,q)] & f &= L+1, L+2, \ldots, \\
x_{n,b}^{on,1}(f-1, q-1) &= \sigma^{LBL-T_S}[x^{on,0}(f,q)] & f &= L+1, L+2, \ldots .
\end{aligned}
\tag{12}
$$

Further, cells arrive at time $k = i \cdot d + 1,\ i = 1, 2, \ldots$ from the beginning of the on-phase. For this arrival instants we obtain the following equations for the probability of the system state:

$$\begin{aligned}
x_e^{on,k}(f-1+I, q-1+T_S) &= \sigma_{QL-T_S}[x_e^{on,k-1}(f,q)], \\
x_{e,b}^{on,k}(f-1+I, q-1) &= \sigma^{QL-T_S}[x_e^{on,k-1}(f,q)] + \pi_{2,1}[x_{e,b}^{on,k-1}(f,q)], \\
x_n^{on,k}(f-1, q-1+T_S) &= \pi_{1,1}[\sigma_{QL-T_S}[x_n^{on,k-1}(f,q)]], \\
x_{n,b}^{on,k}(f-1, q-1) &= \pi_{1,1}[\pi_{2,1}[x_{n,b}^{on,k-1}(f,q)] + \sigma^{QL-T_S}[x_n^{on,k-1}(f,q)]].
\end{aligned} \tag{13}$$

If no cell arrives in a time slot, the state of the F-GCRA and the buffer is decreased by 1. Thus, the state probabilities for $k \neq i \cdot d + 1,\ i = 0, 1, \ldots$ are computed as follows:

$$\begin{aligned}
x_e^{on,k}(f-1, q-1) &= \pi_{1,1}[\pi_{2,1}[x_e^{on,k-1}(f,q)]], \\
x_{e,b}^{on,k}(f-1, q-1) &= \pi_{1,1}[\pi_{2,1}[x_{e,b}^{on,k-1}(f,q)]], \\
x_n^{on,k}(f-1, q-1) &= \pi_{1,1}[\pi_{2,1}[x_n^{on,k-1}(f,q)]], \\
x_{n,b}^{on,k}(f-1, q-1) &= \pi_{1,1}[\pi_{2,1}[x_{n,b}^{on,k-1}(f,q)]].
\end{aligned} \tag{14}$$

The state distribution at the beginning of the off-phase is derived by adding the weighted semidistributions of all possible lengths of the on-phases:

$$\begin{aligned}
x^{off,0}(f,q) = {} & a(0) \cdot x^{on,0}(f,q) + \\
& \sum_{k=1}^{\infty} a(k) \cdot [x_e^{on,k}(f,q) + x_{e,b}^{on,k}(f,q) + x_n^{on,k}(f,q) + x_{n,b}^{on,k}(f,q)].
\end{aligned} \tag{15}$$

Analogously to the case of no arriving cell during the on-phase the system state is decremented by 1 in every time slot of the off-phase. The state probabilities are given recursively by the following equations.

$$x^{off,k}(f-1, q-1) = \pi_{1,1}[\pi_{2,1}[x^{off,k-1}(f,q)]]. \tag{16}$$

Thus, the probability mass function for the state at the beginning of the on-phase computes as follows:

$$x^{on,0}(f,q) = \sum_{k=0}^{\infty} b(k) \cdot x^{off,k}(f,q). \tag{17}$$

We obtain the probability mass function of the system state in equilibrium applying equations (12) to (17) iteratively until convergence is reached.

The probability of a frame being eligible for guaranteed service is derived from the state distribution at the beginning of the on-phase $x^{on,0}(f, q)$. Frames are eligible when the state of the F-GCRA is less or equal to *L*. Thus, the probability that a frame is eligible p_e is given by the following equation:

$$p_e = \sum_{f=0}^{L} \sum_{q=0}^{\infty} x^{on,0}(f, q) . \tag{18}$$

A frame is considered to be lost if at least one cell of the frame is discarded. Thus, the probability of loss for an eligible frame resp. non-eligible frame of length *k* is given by the sum of all values of the vector $x_{e,b}^{on,k}(f, q)$ resp. $x_{n,b}^{on,k}(f, q)$.

Consequently the loss probability of eligible frames $p_{e,f}$ resp. non-eligible frames $p_{n,f}$ computes according to the following expression:

$$\begin{aligned} p_{e,f} &= \sum_{k=1}^{\infty} \sum_{f=0}^{\infty} \sum_{q=0}^{\infty} a(k) \cdot x_{e,b}^{on,k}(f, q), \\ p_{n,f} &= \sum_{k=1}^{\infty} \sum_{f=0}^{\infty} \sum_{q=0}^{\infty} a(k) \cdot x_{n,b}^{on,k}(f, q). \end{aligned} \tag{19}$$

4 Numerical Examples

For the presentation and discussion of numerical examples we will refer to the following basic parameter set unless otherwise expressed.

The length of the on- and off-phase of the considered traffic stream is distributed geometrically. In order to represent a cell indicating the beginning and the end of a frame, the minimum length of an on-phase is set to 2. The system load is chosen to be 25% of the *PCR*, for example for *d=1* – the cells of a frame are sent back to back – the on-phase is set to be 10 slots and the off-phase to 30 slots in average. For other values of *d* the duration of the on- and off-phase is adopted accordingly. The *MCR* of the F-GCRA is set to 20% of the link cell rate, that is *I=5.*

First we look at the queuing behavior of a single link carrying GFR-traffic. In order to reflect the best-effort characteristics of the GFR service the bandwidth available for the connection is set to 1/3 of the *PCR*, that should fulfill the bandwidth requirement of the reference traffic stream. Figure 6 shows the conditional frame loss probability for eligible and non-eligible frames in dependence of the *LBO*. Lower *LBO*

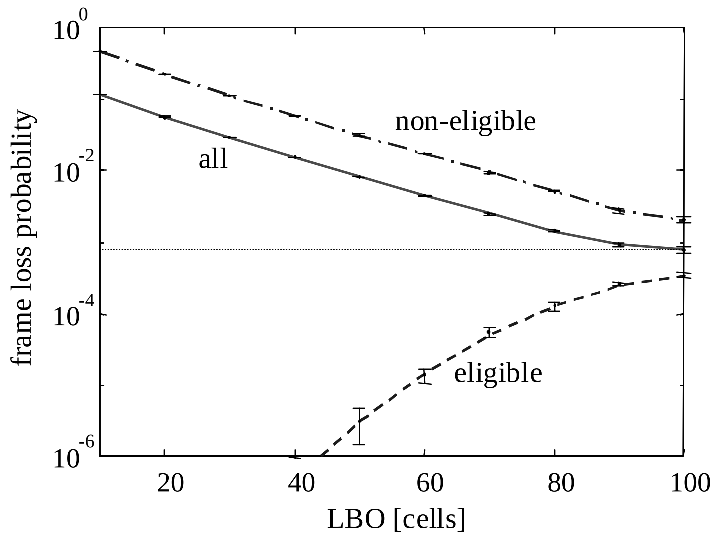

Figure 6: Functionality of the GFR service.

values mean that even in the case of low buffer occupancy non-eligible frames are discarded. Thus, eligible frames are served with higher reliability and suffer lower loss. But on the other hand the transmission buffer for the whole traffic stream is reduced by preferential treatment of the eligible frames, which leads to a reduction of the total throughput of the system.

In Figure 6 we demonstrated that the functionality of the GFR service depends on the proper dimensioning of the *LBO*. To demonstrate the accuracy of our analysis some simulation results are included in the graphic. All analytical values are found to be within the 95% confidence intervals of the simulated reference points.

The basic algorithm (cf. section 2.2) provides a second threshold value *HBO*, which causes the discarding of eligible frames if the buffer is occupied by more than *HBO* cells. For investigating the influence of the *HBO* we used a constant threshold *LBO = 50* in the above mentioned configuration. Since, the value of the *HBO* has to be greater than the value *LBO* to provide a preferred handling of eligible frames the *HBO* was varied between *LBO* and the size of the buffer *Q_MAX*. The results of this investigation are shown in Figure 7. While the loss probability of non-eligible frames is not affected by changing the *HBO*, the loss probability of eligible frames is reduced increasing the *HBO*. If the *HBO* is set to values in the range of the buffer size the loss probability of eligible frames keeps constant. Generally, increasing the *HBO* increases the acceptance of eligible-frames and thus reduces the blocking probability. For large values of the *HBO* the acceptance of eligible frames is also increased, but those frames suffer a higher loss probability due to buffer overflow. To exploit the full range of variation of the *LBO* we set the *HBO* = *Q_MAX* for the remaining numerical examples.

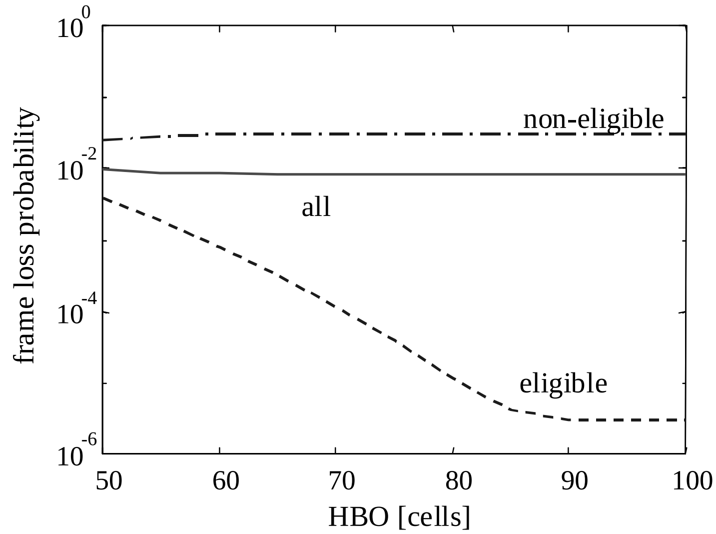

Figure 7: Influence of the *HBO*.

In Figure 8 the blocking probability of eligible frames for different amounts of available bandwidth is depicted. In our example the traffic stream causes a load of 25%. The ratio of eligible frames is 74%. Thus, the system is capable to transfer all eligible frames with a available bandwidth of 20% ($T_S = 5$) of the PCR if the *LBO* is dimensioned properly. If higher capacity is available for the GFR-service, the *LBO* can be chosen higher to obtain the same maximum blocking probability for eligible frames. As indicated in Figure 6 the selection of the highest possible *LBO* leads to the best total throughput.

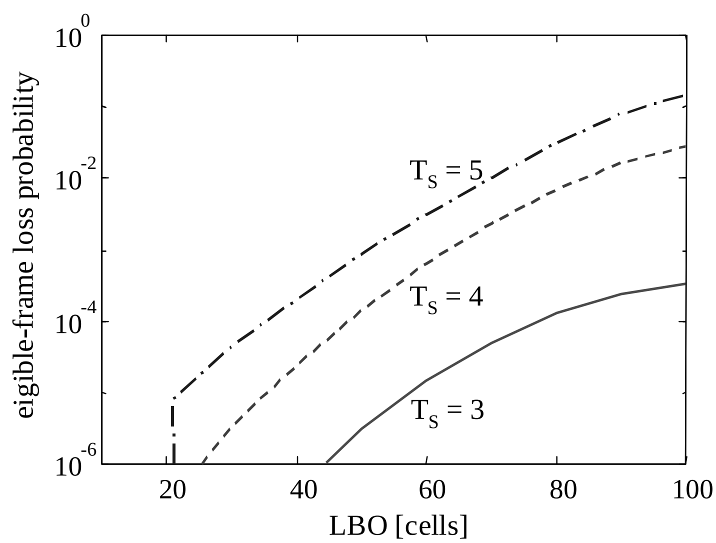

Figure 8: Influence of the available capacity.

To evaluate the impact of the burstiness of the traffic stream on the GFR-service the following system configuration is used: the bandwidth available for the GFR-service is set to 25% of the PCR, that is equal to the minimum bandwidth required by the traffic source. The cells in a frame are spaced by *d* time slots, cf. Section 3.1. The duration of the on- and off-phase are adjusted accordingly to ensure a traffic load of 25%. The parameters of the F-GCRA are adjusted to classify 74% of the frames as eligible, utilizing the analysis introduced in [6]. As shown in Figure 9 the blocking probability of eligible frames increases with increasing burstiness. Again, to obtain an identical maximum blocking probability the parameter *LBO* has to be adjusted accordingly.

The numerical examples show that the dimensioning of the parameters for the GFR-service depends heavily on the traffic characteristics and amount of available bandwidth. Information about the traffic characteristics can be gained from the attributes of a GFR-connection [7] and taken into account. To obtain a high throughput of eligible- and non-eligible frames if bandwidth is available, the – with regard to the service guarantee – highest possible *LBO* value should be chosen. But if the only the minimum bandwidth MCR for the GFR-connection is available the *LBO* has to be chosen more restrictively. A restrictive selection of the *LBO* will fulfill the service requirements in any case, but not the expectations to a best-effort service class. Thus, further investigations should aim at this discrepancy.

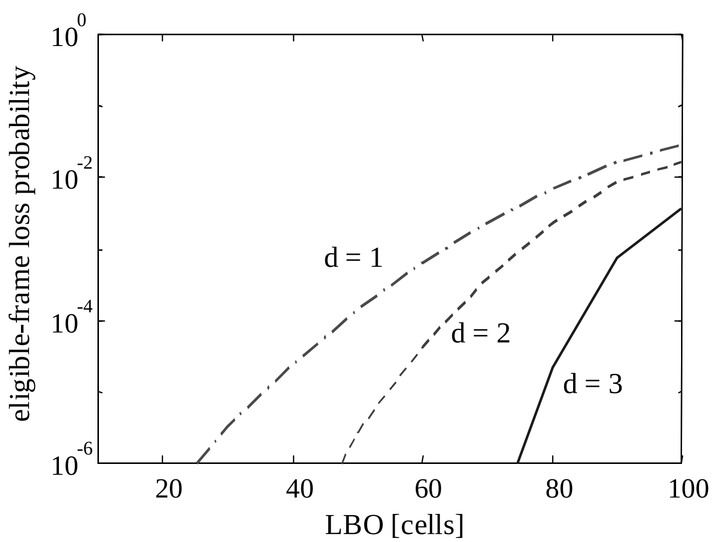

Figure 9: Impact of the traffic-streams burstiness.

5 Summary

In this paper we presented a discrete-time analysis of the GFR-service category, which is currently defined by the ATM Forum. The both key components of the GFR-service – the F-GCRA algorithm and the transmission queue – are described and modeled. While the F-GCRA discriminates eligible and non-eligible frames the buffer discipline ensures the with the GFR-service guaranteed transmission quality. To model the frame-based cell arrivals a on/off-process with generally distributed on- and of-phase was chosen. Since the state of the F-GCRA and the queue are correlated, a two-dimensional discrete-time analysis approach is applied. The numerical examples indicate a discrepancy in the dimensioning of the parameters of the queuing discipline. A restrictive selection of the relevant parameter *LBO* ensures a transmission of eligible frames with the guaranteed service quality, but the best-effort spirit of the GFR-service is given away. Thus, a dimensioning of the queuing discipline of the transmission buffer in dependence of the currently available bandwidth could be an interesting approach to guarantee the service and to preserve the best-effort character of the GFR-service. Further studies will aim at the investigation of the GFR queuing performance under varying load conditions.

Acknowledgment

The financial support of the Deutsche Telekom AG (Technologiezentrum Darmstadt) is appreciated.

References

1. J. Heinanen. *MCR for UBR*. The ATM Forum Technical Committee 96-0362, April 1996.
2. R. Guerin and J. Heinanen. *UBR+ Service Category Definition*. The ATM Forum Technical Committee 96-1598, December 1996.
3. S. Jagannath, N. Yin, J. B. Kenney, J. Heinanen, J. Axell and K. K. Ramakrishnan. *Modified Text for Guaranteed Frame Rate Service Definition.* The ATM Forum Technical Committee 97-0833, December 1997.
4. *Traffic Management Specification Version 4.0*. The ATM Forum Technical Committee, April 1996.
5. *Traffic Management Baseline Text Document, BTD-TM-01-02*. The ATM Forum Technical Committee, July 1998.
6. M. Ritter. *Discrete Time Modeling of the Frame-Based Generic Cell Rate Algorithm.* University of Würzburg, Research Report Series, No. 190, January 1998.
7. G. Koleyni. *Updating Table 2-1.* The ATM Forum Technical Committee 98-0450, July 1998.