IEEE
COMMUNICATIONS
SURVEYS
The Electronic Magazine of
Original Peer-Reviewed Survey Articles

PROVIDING RATE GUARANTEES FOR INTERNET APPLICATION TRAFFIC ACROSS ATM NETWORKS

I. ANDRIKOPOULOS, A. LIAKOPOULOS, G. PAVLOU, AND Z. SUN
UNIVERSITY OF SURREY

ABSTRACT

The TCP/IP protocol suite is the standard requirement for all applications that need to communicate over the Internet. As TCP/IP applications are unable to specify the QoS parameters needed for most Asynchronous Transfer Mode (ATM) services, they tend to use the unspecified bit rate (UBR) service category when running across ATM networks. The UBR service utilizes any bandwidth that is left unused by the rest of the ATM services. This has led the ATM Forum's Traffic Management Group to define a new service category called guaranteed frame rate (GFR). GFR is intended to provide minimum cell rate guarantees and fair access to excess bandwidth left over from higher-priority services. This article first presents a tutorial overview of GFR and then presents a survey of the research work that has been carried out toward the design and implementation of associated ATM switch mechanisms.

Asynchronous Transfer Mode (ATM) has been chosen as the transport technology for the broadband integrated services digital network (B-ISDN). ATM is a connection-oriented switching technology utilizing statistical multiplexing of fixed-length packets, known as cells. ATM was originally designed as a wide area network (WAN) backbone technology offering traffic integration of almost all communication types including voice, video, and data transfer.

The Transmission Control Protocol/Internet Protocol (TCP/IP) [1] is the most widely used protocol suite in computer communications. TCP [2] provides a reliable, connection-oriented, byte-stream transport service using varying-length segments. IP is a connectionless network protocol using varying length packets [1]. Unfortunately, TCP/IP has no notion of traffic requirements comparable to the negotiation of Quality of Service (QoS) parameters in ATM. Since a large number of different workstations and hosts are interconnected using TCP/IP protocols, ATM-based networks must also be capable of carrying traffic of and supporting those protocols. Thus, solutions using ATM as the underlying transport technology for TCP/IP are of great importance, since the wide variety of TCP/IP applications could then be used without modification.

ATM technology is intended to support a wide range of services. Five different ATM service categories have already been defined. These service categories relate the traffic characteristics and QoS requirements to the network behavior. The constant bit rate (CBR) and real-time variable bit rate (rt-VBR) are intended for real-time applications, while the non-real-time variable bit rate (nrt-VBR), the available bit rate (ABR), and the unspecified bit rate (UBR) are intended for non-real-time applications [3]. However, since TCP/IP applications fail to specify the traffic parameters that are needed by most of the ATM services, they tend to use the UBR service. The latter does not provide any rate, cell loss ratio (CLR), or delay variation guarantees and can thus be characterized as a "best effort" service. As the UBR service does not implement any congestion control mechanisms, ATM cells are discarded whenever switch buffers overflow, leading to poor performance for the application traffic.

The guaranteed frame rate (GFR) service, originally called UBR+ [4], is intended to overcome the performance degradation problems faced by applications using the UBR service. The main motivation behind the introduction of GFR was to retain the simplicity of UBR, while providing an enhanced service to the end systems. The main advantage of GFR over UBR is that it allows a minimum guaranteed bandwidth to be associated with each virtual channel (VC) connection, under the assumption of a maximum packet size. Notable differences from UBR are that it explicitly requires the end systems to transmit frames[1] and also requires the ATM switches to be aware of the frame boundaries. This means that congested ATM switches should normally discard entire frames instead

---

This work was partially supported by the EU IthACI ACTS Project (AC337).

[1] The term frame means an AAL5 protocol data unit (PDU). Both terms are used interchangeably throughout the article.

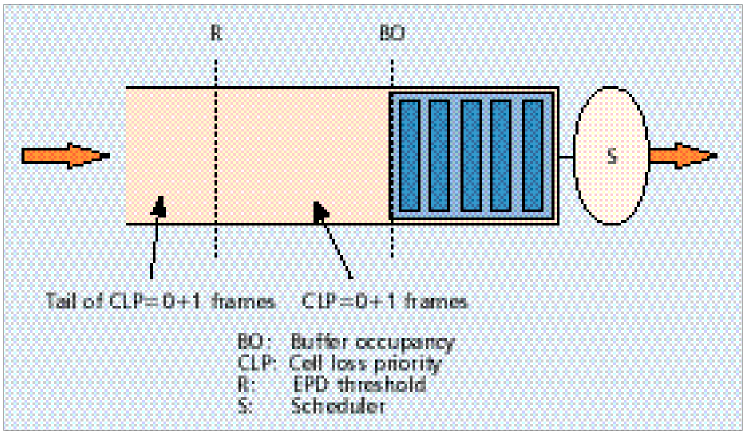


FIGURE 1. The EPD buffer policy.

of discarding individual cells. In case an incomplete frame is delivered, the network should try to deliver the last cell of that frame to indicate the end of the frame. Frames are delivered at the receiving end conforming at least to a minimum cell rate specified in a traffic contract. Any frames sent in excess of the traffic contract are delivered within the limits of the fair share of the available bandwidth left over from higher-priority connections.

The interaction between ATM and the applications running over it can be minimal, leading to a large extent in preserving their simplicity. The primary application suited for use of GFR is LAN interconnection via ATM backbone networks, offering more flexibility than traditional leased lines. The backbone networks can either be private, e.g., an ATM corporate backbone interconnecting several local area networks (LANs), or public, e.g., an ATM backbone interconnecting several Internet service provider (ISP) networks. A GFR VC can be used to transport the traffic of multiple TCP/IP connections between the LANs, offering the possibility of providing virtual private network (VPN) services. Each VC can always transmit at the minimum guaranteed bandwidth, even if the network is congested, and will be able to transmit at a higher rate when the network is not congested. GFR could also be useful to IP routers separated by ATM networks, where GFR VCs can be used to carry aggregated traffic between the routers.

Complex mechanisms need to be implemented inside the ATM switches so that GFR guarantees can be supported. We can separate GFR mechanisms into two groups. The first group includes all the FIFO-based (first-in-first-out) mechanisms that use tagging and buffer management to allocate a portion of the FIFO buffer space to each connection. The second group includes scheduling disciplines for providing rate guarantees. In addition, policing methods are used to isolate the traffic sources from each other.

This article provides a tutorial overview of GFR that is further expanded with a survey on the research work that has been done toward GFR's definition and implementation. A number of representative mechanisms that have been proposed for implementing the GFR service in ATM switches are discussed and compared based on conclusions deduced from simulation results. These mechanisms are a combination of different buffer management and scheduling schemes, together with service conformance and QoS eligibility test mechanisms. There is a wide spectrum of possible GFR implementations in the literature, where each implementation varies in terms of performance and complexity. The most basic of these implementations are covered in this article.

The reader is expected to have an understanding of the basic concepts of both ATM and TCP/IP.

ATM SERVICE CATEGORIES

ATM provides the CBR, rt-VBR, nrt-VBR, ABR, and UBR services in order to support applications with a wide variety of requirements. These services differ significantly in terms of their characteristics and QoS requirements. Precisely defined guarantees for the bit rate, delay, and cell loss are provided by the CBR and rt-VBR services, while less strict requirements are provided by the ABR and nrt-VBR services. The UBR service category does not provide any QoS guarantees.

GFR is a major enhancement to UBR and has been elected as a new ATM service category [5]. Nevertheless, GFR differs considerably from UBR in that it provides minimum bandwidth guarantees, whereas UBR does not.

In the following sections, we discuss the UBR and GFR service categories.

THE UNSPECIFIED BIT RATE SERVICE CATEGORY

The UBR service category does not offer any service guarantees to the end-system applications that use it. These applications are allowed to send any amount of data, but the network does not provide any guarantees for the cell loss ratio or for the delay variation that the traffic might experience. UBR is, therefore, used for non-real-time or elastic applications, such as TCP/IP-based file transfer, which do not require any tightly constrained delay or delay variation.

UBR does not mandate any congestion-control mechanisms. These may be performed by higher layers (e.g., the TCP layer) at the end systems. The ATM switches will discard cells when their buffers overflow. The absence of network-based congestion control can therefore lead to poor performance for the applications.

To improve performance, many discard policies have been proposed. The early packet discard (EPD) policy is generally considered the most efficient. If congestion occurs at the ATM switch, the EPD policy drops complete IP packets instead of individual cells or partial packets, as does the partial packet discard (PPD) policy [6]. As a result, the number of incomplete IP packets[2] that are transported over the network is minimized and better network utilization is consequently achieved. The EPD mechanism uses a static threshold R that is less than the buffer size. When mild congestion occurs, i.e., the buffer occupancy BO exceeds the threshold R, all the cells from newly arriving packets are dropped. The cells that belong to partially received packets are held for switching as long as there is enough space in the buffer, as illustrated in Fig. 1.

The UBR service is used extensively by TCP/IP-based applications. However, the absence of network congestion control mechanisms leads TCP/IP applications to experience low performance and high unfairness when congestion occurs. Simulations have shown that EPD can improve the efficiency of TCP but not its fairness to different connections [7].

THE GUARANTEED FRAME RATE SERVICE CATEGORY

The GFR service [5, 8] is designed to provide, with high probability, packet-based or frame-based applications (e.g., TCP traffic) with a minimum guaranteed cell rate through an ATM network. A key assumption is that the frame length

---

[2] Each IP packet corresponds to one AAL5 PDU/frame.

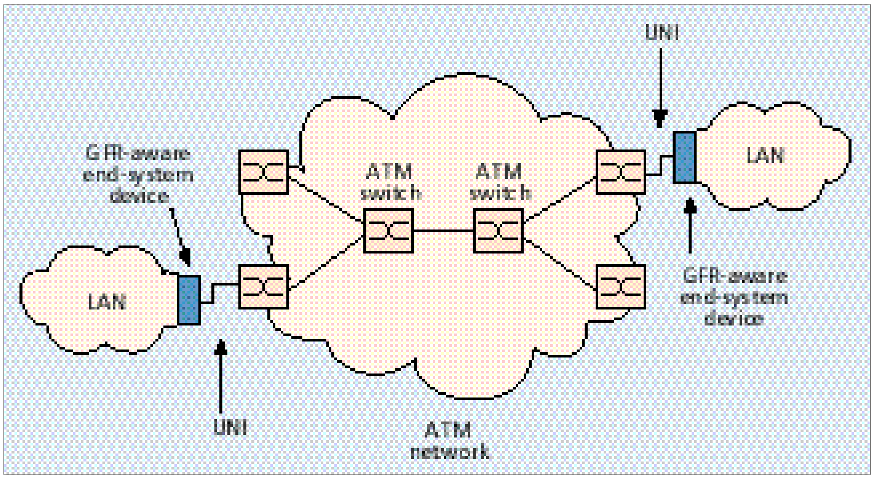


FIGURE 2. Location of "GFR-aware" edge devices.

does not exceed a maximum frame size (MFS) in a burst that does not exceed a maximum burst size (MBS) [5, 8]. The GFR service also allows the user to send traffic in excess of the agreed minimum cell rate (MCR) and the associated MBS, but the excess traffic will only be delivered within the limit of available resources.

GFR service guarantees are based around AAL5 frames, and under congestion the network aims to discard complete frames, a function known as frame discard [5], instead of discarding arbitrary cells. However, in case an incomplete frame is delivered by the network, the last cell of that frame should also be delivered to indicate the end of the frame. GFR only applies to VCs since frame delineation is not generally visible in a virtual path connection (VPC). Frame delineation inside the network takes place at every ATM switch by detecting the AAL5 PDU boundaries. The detection is done by examining a particular bit in the payload type (PT) field of each cell header. All the cells of the same frame have this bit set to 0 except from the last cell, which has its bit set to 1 to indicate the end of the frame.

Apart from the minimum cell rate guarantees, VCs using the GFR service are expected to be able to fairly use any additional bandwidth left over from higher-priority services. In other words, if the end system sends frames at a cell rate less than or equal to the minimum guaranteed cell rate, then all the frames are expected to be delivered at the receiving end with minimum loss. If the end system sends frames at a cell rate higher than the guaranteed cell rate, then the receiving end should receive at least the minimum cell rate. Excess traffic will be delivered within the limits of available resources in a best-effort manner. Moreover, the service specifies that the excess traffic from each user should receive a fair share of the unused network bandwidth. From the network point of view, supporting the GFR service demands that the network must be capable of supporting a minimum cell rate for all existing connections and under all possible conditions. The network should treat identical connections, i.e., connections with the same QoS requirements and traffic contract, in the same way and provide similar final service (fairness).

As with the other ATM service categories, a traffic contract must be specified for a GFR connection. The GFR traffic contract is composed of four parameters:

- Minimum cell rate (MCR) and associated cell delay variation tolerance ($CDVT_{MCR}$): the minimum cell rate agreed between the end systems and the network for a specific ATM connection.
- Peak cell rate (PCR) and associated $CDVT_{PCR}$: the upper bound on the cell rate, at which traffic can be submitted on an ATM VC.
- Maximum frame size (MFS): the maximum AAL5 frame size in cells.
- Maximum burst size (MBS): the maximum number of cells that may arrive at a rate equal to the PCR.

The CDVT is the measure of departure from exact periodicity of cell arrival on a specific ATM VC.

The MCR is the rate at which cells are produced by the segmentation of frames and it must be strictly less than the PCR. The size of the frames should be at most equal to the MFS in order to be considered for GFR guarantees, and the MBS should be strictly greater than the MFS. The MFS is actually related to the protocols used above the AAL5 layer. For example, the default maximum transmission unit (MTU) for classical IP over ATM is 9188 bytes, or 192 cells including the 28 bytes of AAL5 overhead [9], indicating an MFS of the same size.

The PCR, MCR, MBS, and MFS parameters, together with the CDVT, are defined at the time of subscription of a permanent connection or, in the case of switched connections, they are carried in an ATM traffic descriptor field of the SETUP message at connection control, via ATM signalling [10]. The calling user has the options to negotiate the above parameters with the network, to use default values, and even to define minimum acceptable values that can be used by the network during the call establishment procedures in the case where the requested traffic contract cannot be satisfied. The CDVT may be chosen by the network, and it takes into account any perturbation that may affect the conformance of the traffic to its traffic contract, e.g., cell multiplexing occurring at end systems.

The requested GFR guarantees are associated directly with the MCR, MBS, and MFS parameters, which can be very easily determined provided the traffic source characteristics are known. However, the end system needs to be aware of the existence of the underlying ATM network in order to request an ATM VC connection with GFR capability. It is likely that a GFR service user may not be able to specify the traffic source characteristics that are necessary to define the above parameters. In that case, using default values for the MCR and MBS, or choosing, based on heuristics, the values of the basic parameters corresponding to the traffic source characteristics, is an important research issue [11, 12]. When the GFR service is used particularly for LAN or router interconnection via an ATM infrastructure, then the LAN edge devices at the user-network interface (UNI) is the only equipment that needs to "understand" GFR. This is illustrated in Fig. 2.

METHODS FOR SUPPORTING RATE GUARANTEES

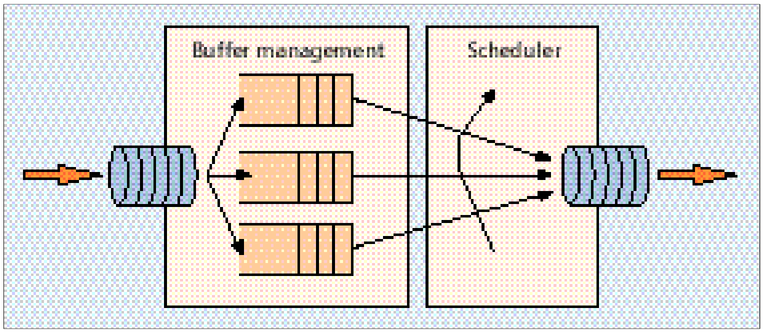
Generally, there are three basic approaches that can be used to provide network-level QoS to applications and to enable a number of users to efficiently use and fairly share the available network bandwidth. These are:

- Policing
- Buffer management
- Scheduling

Various combinations of these methods can lead to a wide spectrum of possible GFR implementations. In the following subsections, we will briefly describe these methods.

POLICING

The policing method is used for tagging non-conforming cells so that the conforming and non-conforming traffic can be distinguished. Tagging is done by converting the CLP bit of the ATM cell header from 0 to 1. As the tagged cells are assumed to have violated the traffic contract, they are not considered eligible for rate guarantees. Therefore, subsequent mechanisms, such as buffer management, will forward untagged cells in preference to tagged cells. Tagging should be performed at the frame level, i.e., all the cells of a frame should be identically tagged.



■ **FIGURE 3.** Typical arrangement of the buffer management and scheduling mechanisms.

Tagging is generally performed in the network and must be signalled during connection establishment. However, tagging can be performed by the source end system in order to indicate less important frames, in which case it is called marking. This scheme can provide higher priority to frames that carry more important information. Alternatively, tagging may be performed at the ingress of the network. It should be noted that when a network element receives a tagged cell, it cannot distinguish whether the cell has been tagged by the end system or by the network.

Finally, tagging methods take into account the traffic eligibility criteria, as well the network congestion, in order to improve the fairness in the allocation of the unused bandwidth.

BUFFER MANAGEMENT

The purpose of buffer management schemes is to control the number of cells stored in the network element buffer. These methods aim to satisfy the cell loss requirements of the different services, and to utilize the buffer as efficiently as possible. Therefore, the buffer management mechanisms decide whether an arriving cell should be stored or discarded.

In the event of congestion, the buffer management methods always discard the tagged cells in preference to the untagged cells. Schemes that are more complex may push low-priority cells out of the buffer in order to free space for storing high-priority cells.

The buffer management schemes may organize the switch buffer into a number of logical buffers, each of which is associated with a specific connection (per-VC queuing). In addition, per-VC accounting may be used to control the buffer space that each VC occupies. Per-VC accounting consists of maintaining a per-VC variable, which reports the current queue occupancy of each active flow sharing the output buffer. Hence, different connections can be handled differently. Cell discarding can be based on queue-specific occupancy thresholds. These schemes can potentially provide the needed traffic isolation, but they are more difficult to implement in the network elements.

SCHEDULING

Scheduling mechanisms determine when and in which order per-VC queues should be served. The scheduler can control the outgoing rate of individual VCs and thus ensure that the individual connections use a fair portion of the available bandwidth. The scheduling schemes usually control only the bandwidth that one connection occupies in a link, and do not provide explicit delay guarantees to the traffic. However, in most of these cases, the rate guarantees result in end-to-end delay guarantees. More complex implementations can provide explicit guarantees in terms of minimum cell rates and maximum end-to-end delays [13].

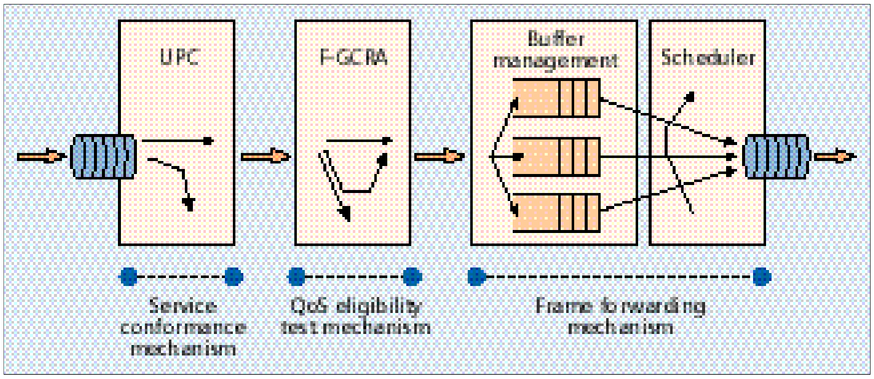
The scheduling schemes can be implemented together with buffer management mechanisms, as illustrated in Fig. 3.

How the above methods can be used to provide the GFR capability is discussed in the next section.

MECHANISMS FOR PROVIDING GFR GUARANTEES

Minimum rate guarantees are provided by a set of three mechanisms that independently control the cell flow. These mechanisms are based on the design methods mentioned previously and usually implement a combination of these methods. In other words, a mechanism may use specific policing, buffer management, and scheduling implementations to provide minimum rate guarantees.

In the following subsections, we describe these mechanisms according to the sequence in which they may be implemented. Figure 4 shows a graphical representation of these mechanisms.



■ **FIGURE 4.** The fundamental components of a GFR mechanism.

SERVICE CONFORMANCE MECHANISM

A first step in providing QoS, such as providing basic rate guarantees, is to ensure that traffic from a misbehaving connection will not affect the other already established connections. The usage parameter control (UPC) function is responsible for this task. The UPC is a set of actions that constantly monitor the network traffic from all the active VCs and ensure that the traffic of the connections remains conforming with the negotiated traffic contract. These actions deal with checking the validity of the virtual path identifier (VPI) and the virtual channel identifier (VCI), and monitoring the traffic entering the network from active[3] VCs or virtual path connections (VPCs) in order to ensure that the parameters agreed upon are not violated. In case of traffic violation, i.e., cells sent in excess of the traffic contract, the UPC mechanism may tag or discard the non-conforming cells. The UPC is usually performed at the UNI. When performed at the network node interface (NNI) it is called network parameter control (NPC) [5].

As was previously outlined, GFR guarantees a minimum cell rate service to frames that do not exceed the maximum frame size. Therefore, the GFR UPC performs the following roles:

- Verifying that the frame size is not greater than the maximum allowed frame size.
- Verifying that the end systems send traffic that conforms to the PCR and $CDVT_{PCR}$ traffic parameters.
- Verifying that all the cells of a frame are tagged identically.

The UPC mechanism may tag or discard cells that belong to frames sent in excess of the PCR parameter, i.e., non-conforming frames. Note that, according to the GFR definition, all the cells belonging to the same frame must be tagged identically. However, when the first cell of the frame is received, the UPC mechanism is unable to predict whether all the succeeding cells will be conforming or not. Thus, the cells of non-conforming frames are usually discarded. Additionally, the UPC discards the tail cells of frames that exceed the maximum frame size.

The GFR UPC may be based on the generic cell rate algorithm GCRA (T, τ). The GCRA is used to define conformance with respect to the traffic contract that controls the maximum cell rate. The GCRA has two parameters, T and τ, and it times the arrivals of cells as follows.

The theoretical arrival time of a cell is defined as TAT. If the next cell arrives before TAT – τ, then the algorithm GCRA (T, τ) declares that cell to be non-conforming, and TAT is unchanged. If it arrives at time $t \geq TAT - \tau$, then the cell is conforming and the algorithm resets the value of TAT to max{t, TAT} + T. Figure 5 illustrates the three possible cell arrival events.

To restrict the maximum cell rate to the PCR and its associated $CDVT_{PCR}$ the GCRA(1/PCR, $CDVT_{PCR}$) algorithm is used.

Additional tests verify the conformance of the traffic with the MFS and with the cell tagging requirements. We will not elaborate more on the GFR UPC mechanism or its implementation. However, representative pseudo-code can be found in [14].

QOS ELIGIBILITY TEST MECHANISM

The frame-based generic cell rate algorithm (F-GCRA) is a modification of its widely known GCRA counterpart, and is used by the network to identify the conforming frames that should be eligible for service guarantees. More accurately, the F-GCRA(T, f) mechanism is determined by the parameters T = 1/MCR and tolerance $f \geq BT + CDVT_{MCR}$ where the burst tolerance is

$$BT = (MBS - 1)\left(\frac{1}{MCR} - \frac{1}{PCR}\right).$$

The T value is related to the MCR parameter of the GFR traffic contract, while the tolerance f takes into account the bursty characteristics of the traffic and the delay variations in the network. Additionally, the PCR is strictly greater than the MCR and

$$MBS \geq 1 + \frac{MFS\ PCR}{PCR - MCR}.$$

The MCR should strictly be greater than zero, otherwise in the case where the MCR is set equal to zero, there are no eligible frames and therefore the F-GCRA test does not apply.

The F-GCRA algorithm can be described by the pseudo-code given below [5]:

```
/* Initialization at the arrival of the
first cell at time $t_a$ after the
establishment of the GFR connection. */
X = 0
LCT = $t_a$
passed = false

/* At the arrival of the first cell of a
frame at time $t_a$ */
X’ = X – ($t_a$ – LCT)
if (X’ > f) or (tagged cell)
      /* tag cell */
      passed = false
else
      /* cell is conforming */
      passed = true
endif

if (passed) then
      /* cell is conforming hence
      deposit token in bucket and
      update variables */
      X = max(0,X’) + T
      LCT = $t_a$
endif

/* At the arrival of $i^{th}$ cell
of the frame at time $t_a$ */
if (passed) then
      /* cell is conforming hence
      deposit token in bucket and
      update variables */
      X’ = X – ($t_a$ – LCT)
      X = max(0, X’) + T
      LCT = $t_a$
endif
```

The F-GCRA mechanism consists of a bucket of tokens, whose capacity is given by the leaky bucket counter. The counter is denoted by the non-negative variable X. X is initialized to zero when the connection is established. The bucket drops a token every T = 1/MCR, or in other words, drains out at a constant rate equal to the MCR. At the time of arrival t_a of the first cell of the frame, the variable X’, measuring the guaranteed service granted in excess of the MCR, is calculated according to the last conformance time LCT. If the bucket is filled with tokens up to the limit f, i.e., $X' \leq f$, then the cell is considered conforming. In this case, the auxiliary variable passed is set to true, one token is deposited in the bucket according to X = max(0,X’) + T, and LCT is set to the current value of t_a. Each subsequent cell of the conforming frame updates X’ with respect to its arrival time t_a and also consumes a token from the bucket when the cell is forwarded, thus updating the variable X accordingly. LCT is then set

---

[3] A VC is said to be active if at least one of its cells is residing in the switch buffer.

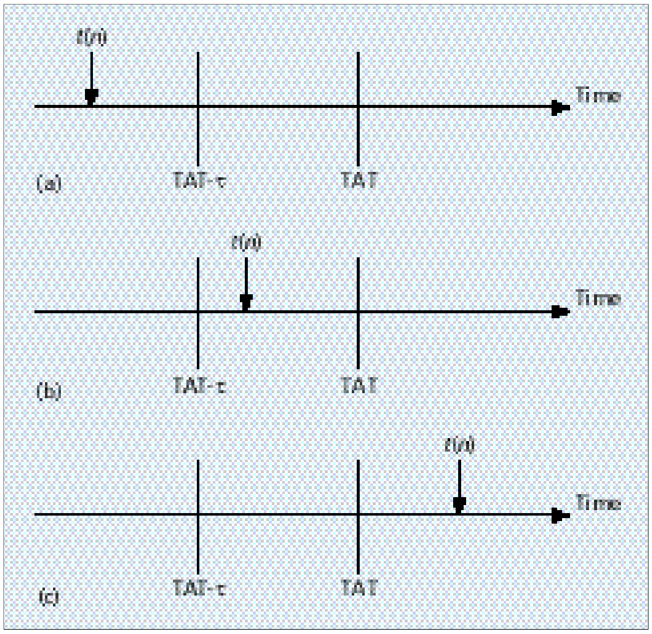


FIGURE 5. (a) nth cell is too early and non-conforming; (b) n^{th} cell is early but conforming; (c) n^{th} cell is conforming.

equal to the new t_a. Otherwise, if at the time of arrival of the first cell of the frame the bucket is filled up with tokens that exceed the limit f, or this cell is tagged, then all the cells of the frame are considered non-conforming, i.e., passed is set to false. These cells may be tagged, if the end system allows network tagging, or discarded. None of these cells consumes a token in the bucket when forwarded. It should be noted that the tagged cells are considered a priori non-eligible for rate guarantees. Thus, the arrival of a tagged frame, i.e., a non-conforming frame, will not increase the number of tokens in the bucket.

The algorithm described above is the one proposed by the ATM Forum [5]. However, at this point it is interesting, for comparison purposes and as background information, to mention the QoS eligibility test mechanism that was initially proposed [4]. This mechanism was widely used in the simulations presented in relevant published papers and was also based on a modification of the GCRA mechanism. Like GCRA, it consisted of a bucket that was filled with tokens at an MCR rate. If at the time of arrival of the first cell of a frame the number of tokens in the bucket was greater than MBS/2, then all the cells of the frame were considered eligible. Each of the cells, when it was forwarded, consumed a token from the bucket. Otherwise, if at the time of arrival of the first cell of a frame there were fewer than MBS/2 tokens in the bucket, the frame was considered non-eligible. The cells of this frame could be tagged or discarded, and none of them consumed a token from the bucket when they were forwarded. The value of MBS/2 corresponded to one frame measured in cell units, as the BT parameter was set to twice the frame size. The pseudo-code for these implementations can be found in [15].

Depending on whether network-based tagging is allowed, two conformance definitions have been defined for GFR, namely GFR.1 and GFR.2.

Under the GFR.1 conformance definition, the GFR service user may request that tagging is not performed by the network. This means that it is up to each ATM switch along the network path to choose which CLP = 0 frames will be eligible for minimum cell rate guarantees and which could be discarded (non-eligible frames) as part of the excess traffic, if congestion occurs.

On the other hand, GFR.2 allows network tagging and therefore the F-GCRA test may be used by the network to tag non-eligible frames. The ATM switches in the network can rely on the CLP bit to distinguish between eligible CLP = 0 and non-eligible CLP = 1 frames so as to provide the necessary minimum guarantees. Non-eligible traffic will be delivered by the network on a best-effort basis.

It is anticipated that GFR.1, in contrast to GFR.2, will require implementation of sophisticated mechanisms in every ATM switch along the network path to inhibit CLP = 1 traffic from affecting the provision of MCR guarantees to CLP = 0 traffic. On the other hand, if network-based tagging, i.e., GFR.2, is allowed, it is possible that it can influence to some extent the level of the GFR service being provided, if tagging decisions are exclusively based on the F-GCRA test without taking into consideration congestion conditions. Fair allocation of the available bandwidth to excess traffic could then be affected.

Finally, it is important to clarify that a network that supports GFR guarantees does not provide rate services precisely to the frames/cells that are considered conforming by the F-GCRA mechanism. The network has to provide the supported service for at least a number of cells equal to the number of conforming cells according to the F-GCRA mechanism [5]. Consequently, some of the conforming frames occasionally may not receive the GFR service guarantees.

FRAME FORWARDING MECHANISM

The frame forwarding mechanism is the most complex part of a GFR mechanism. It is composed of two main parts that operate in sequence. First, the arriving cells are stored into the network element buffers according to the information provided by previously applied policy mechanisms, such as F-GCRA. Second, the buffered cells are forwarded to the network.

A frame forwarding mechanism can be characterized by its buffering, e.g., per-VC queuing, and its scheduling, e.g., weighted round robin (WRR), schemes. Usually, the more complex these schemes are, the more effectively the rate guarantees can be supported. On the other hand, it should be pointed out that GFR mechanisms should require minimal modifications to the network elements. Therefore, simple mechanisms are always preferable.

Finally, it should be noted that the frame forwarding mechanisms are often inaccurately considered as mechanisms that provide GFR service guarantees. This is a misconception, since without the service conformance and the QoS eligibility test mechanisms a frame forwarding mechanism is unable alone to sustain the GFR services.

SWITCH DESIGN OPTIONS FOR SUPPORTING GFR GUARANTEES

This section gives an overview of a number of representative frame forwarding mechanisms that have been studied and have been proposed for implementation in ATM switches in order to support the GFR service category. Also presented are descriptions of the design of the different mechanisms together with simulation results and comments about whether it is possible to provide minimum bandwidth guarantees.

Since TCP/IP application traffic is the target traffic type to be conveyed over GFR VCs, simulations have been carried out to evaluate the performance of TCP over GFR. The basic conclusions from these simulations are provided as they indicate which forwarding mechanisms deserve to be considered for a GFR implementation. It should be pointed out that we

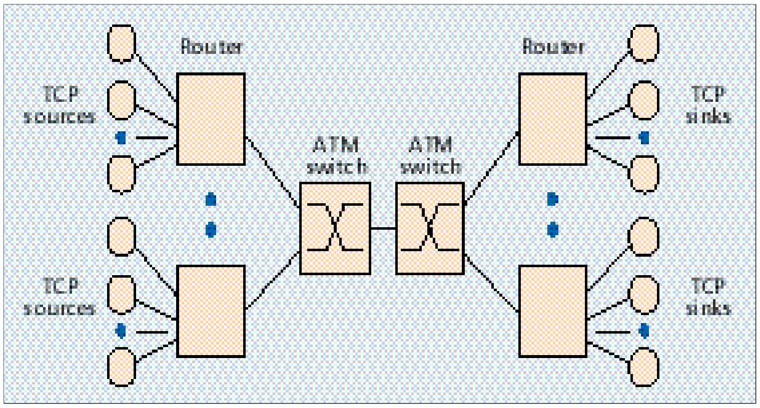


FIGURE 6. A typical network configuration used in the simulations.

make an attempt to summarize the most important conclusions deduced from the referenced simulations without going into extreme details that can otherwise confuse the reader. A more detailed analysis is beyond the scope of this article. The reader can consult the papers and contributions listed in the references to find more detailed information about the simulations that were used.

The reader should pay attention to the fact that GFR does not guarantee TCP throughput. However, guaranteed bandwidth at the frame level may yield TCP throughput performance that is proportional to the guaranteed minimum cell rate. This is something that depends on the GFR service conformance definitions used. It also depends on the right choice of the GFR traffic contract parameters so that they reflect the source TCP traffic characteristics as accurately as possible.

For a given lossless TCP connection, the maximum TCP throughput that can be achieved is proportional to the congestion window and to the inverse of the connection round-trip time. How the congestion window changes over time depends on the current congestion control phase which in turn depends mainly on the packet loss that may occur during the duration of the connection, as well as the round-trip time [16, 17].

If we now want to guarantee TCP throughput by requesting the GFR service with a specific MCR, we must allow the TCP congestion window to increase up to the value W that corresponds to the requested MCR. Any TCP “frames” sent with a window size larger than W will be tagged and delivered on a best-effort basis. If congestion occurs the tagged frames will be discarded. Considering the fact that the maximum allowed window size corresponding to MCR is W, MBS can be set to W · m, where m is the frame size in ATM cells [11]. It should be noted, however, that this sizing rule is a simplistic heuristic formula rather than a proven rule with evident accuracy.

SIMULATION SCENARIOS

Due to the nature of this article, it is not possible to describe all the simulation scenarios that are mentioned in this article. Most simulation scenarios are based around similar network topologies such as the one illustrated in Fig. 6. Each paper utilizes a different set of network components and parameters (e.g., number of TCP sources and sinks, number of routers, discard thresholds, link delay, link capacity, etc.). TCP characteristics, such as TCP version, maximum window size, retransmission mechanism, timer granularity, use of delayed acknowledgements, etc., also differed among the papers. The TCP versions, whose performance was evaluated, were TCP Tahoe, Reno, and SACK [18].

Consequently, these differences in the simulation configuration among the various papers make it difficult to draw general conclusions about the effectiveness of the stated mechanisms. The reader, though, should not forget that the main target application for GFR is LAN interconnection, and hence the simulation scenarios corresponding to such an application should normally be selected for the correct evaluation of the mechanisms. Not all contributions and papers use this approach, however.

FIFO-BASED WITH FRAME TAGGING

The simplest mechanism that has been proposed is a FIFO-based mechanism with policing (tagging). It consists of a single buffer logically divided into three parts with the use of two thresholds, namely the low buffer occupancy (LBO) and the high buffer occupancy (HBO), as shown in Fig. 7 [4].

The cell drop policy, which is known as double-EPD, depends on the buffer occupancy (BO) variable that determines the maximum number of cells residing in the buffer. When the buffer is unloaded, i.e., the BO is below the LBO threshold, all the cells are queued. As the BO exceeds the LBO, yet still remains below the HBO threshold, all the cells belonging to newly arriving tagged frames are discarded. Untagged and tagged cells belonging to partially accepted frames are stored. The tagged cells are always discarded in preference to the untagged cells. The cell drop decisions for the first cell of a frame and for subsequent cells of the same frame for an example implementation are summarized in Tables 1 and 2, respectively.

Finally, as the BO exceeds the HBO threshold, only the untagged cells belonging to the partially accepted packets are stored. The HBO threshold is identical to the classical EPD threshold. In summary, as the BO exceeds the LBO and HBO thresholds, the cell drop policy becomes gradually stricter and cells are likely to be discarded [19].

Pappu and Basak argued that a FIFO-based mechanism with policing is insufficient to provide GFR guarantees. Using VCs with a single TCP connection per VC, they found in their simulations that sources with high MCR failed to make use of their proportion of the bandwidth [15].

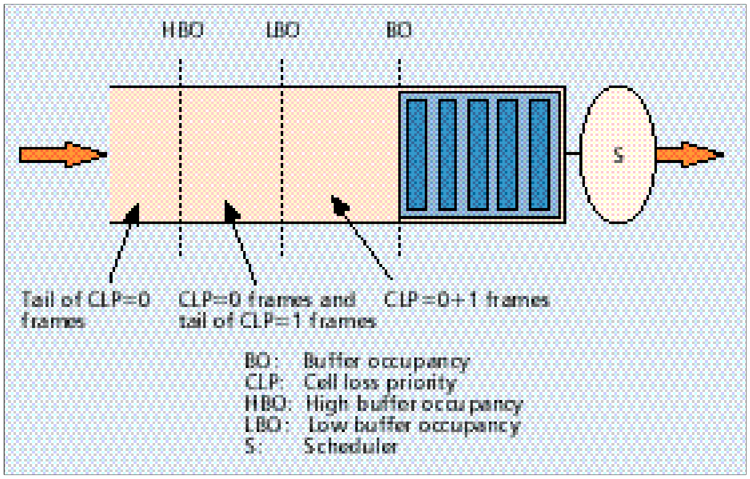


FIGURE 7. The FIFO-based with frame tagging mechanism.

Nonetheless, Bonaventure in [19] found that when several TCP connections are multiplexed by a router on a single VC — which actually forms a more realistic scenario than the previous one — the MCR can be utilized. In addition, the sources attached to each router were much influenced by the MBS. With large MBS, e.g., 20 times the MFS, the sources could utilize their reserved bandwidth and the influence of the router-to-router delay was limited. In [14] the influence of the TCP congestion control mechanism on TCP performance was studied. Among the various different versions of TCP that were examined with the particular GFR forwarding mechanism analyzed here, the SACK TCP [20] demonstrated the best performance, although the results were considered far from being satisfactory. It was also shown that apart from the coarse TCP timer granularity, one of the main factors leading to a severe TCP throughput performance degradation is the use of a modified GCRA at the network ingress to tag the frames for a FIFO-based switch. The reason given for the observed behavior was that TCP had difficulties in adapting its rate to a traffic contract enforced by the F-GCRA. It was, therefore, proposed that if FIFO-based switches are to be maintained, tagging should be done inside the network and not at the network access point. Furthermore, the TCP throughput was not found to be particularly affected by other GFR traffic sharing the link as well as of the threshold HBO. Small values of the LBO produced somewhat better results for a TCP with small granularity timer. Finally, the value of the TCP maximum segment size (MSS) did not have any effect on TCP throughput, provided the MSS was the same for all TCP sources. Different MSS values among TCP sources reduced the throughput fairness significantly, as the sources with larger MSS achieved better throughput performance [14].

FIFO-BASED WITH FRAME TAGGING AND PER-VC QUEUING

Its frame forwarding mechanism is an enhancement of the simple FIFO-based counterpart. In this mechanism, the switch buffer is divided into a number of logical queues, where each queue is associated with a specific VC (per-VC queuing), as depicted in Fig. 8.

As in the previous implementation, two thresholds, LBO and HBO, are used to partition the logical buffer into three different parts. Furthermore, per-VC thresholds R_i are set in accordance with each VC's MCR. These thresholds are specified when the VCs are established and they remain static for the duration of the connection. When switch congestion occurs, i.e., the buffer occupancy exceeds the threshold HBO, each connection can store a newly-arriving untagged frame if there are not many of its cells in the buffer. This is illustrated in Fig. 8, where cells from only VC_1, i.e., currently the VC_i for which the queue length is less than R_i, are accepted in the buffer. As shown in Table 3, in case of mild congestion the drop decision for the first cell of an untagged frame is conditional. In [21] it was shown that the obtained TCP throughputs were in proportion to the allocated weights and that the link utilization was always close to 100 percent.

An approach that provides rate guarantees to individual or a set of flows with a FIFO scheduler simply by relying on a simple buffer management scheme has also been proposed and analyzed in [22].

Buffer occupancy	***BO<LBO*** **unloaded buffer**	***LBO<BO<HBO***	***HBO<BO*** **mild congestion**
Untagged	Accept	Accept	Discard
Tagged	Accept	Discard	Discard

Table 1. Cell drop decision at the arrival of a new frame.

Buffer occupancy	***BO<LBO*** **unloaded buffer**	***LBO<BO<HBO***	***HBO<BO*** **mild congestion**
Untagged	Accept	Accept	Accept
Tagged	Accept	Accept	Discard

Table 2. Cell drop decision for subsequent cells.

Buffer occupancy	***BO<LBO*** **unloaded buffer**	***LBO<BO<HBO***	***HBO<BO*** **mild congestion**
Untagged	Accept	Accept	Conditional
Tagged	Accept	Discard	Discard

Table 3. Cell drop decision at the arrival of a new frame (II).

FIFO-BASED WITH FRAME TAGGING AND FBA (PER-VC ACCOUNTING)

This mechanism uses a sophisticated buffer allocation scheme called fair buffer allocation (FBA). FBA tries to divide the buffer occupancy fairly among the active connections. Therefore, the per-VC congestion thresholds change over time as a function of the connection buffer occupancy and of the free buffer space. As the incoming cells are stored in the FIFO buffer, rate guarantees can be achieved only if the buffer is divided in accordance with the per-VC MCR traffic parameters.

Goyal et al. investigated in [7] whether the FIFO-based mechanism would be adequate for providing GFR guarantees if it was enhanced with an FBA scheme presented in [23]. The frame discarding decision was based on the following conditions:

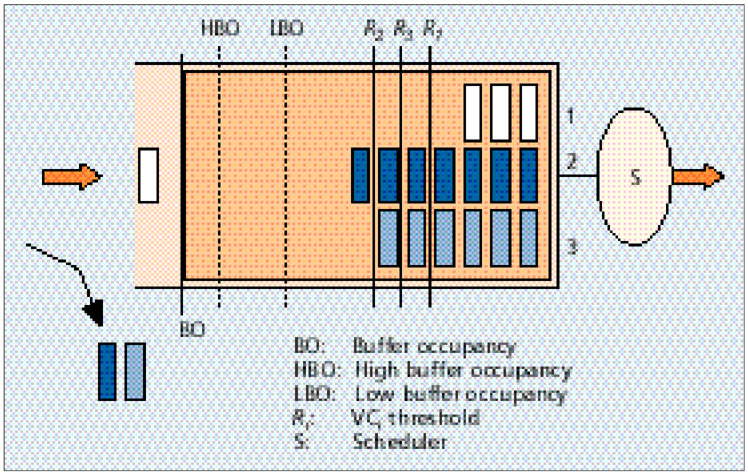


FIGURE 8. The FIFO-based with frame tagging and per-VC queuing mechanism.

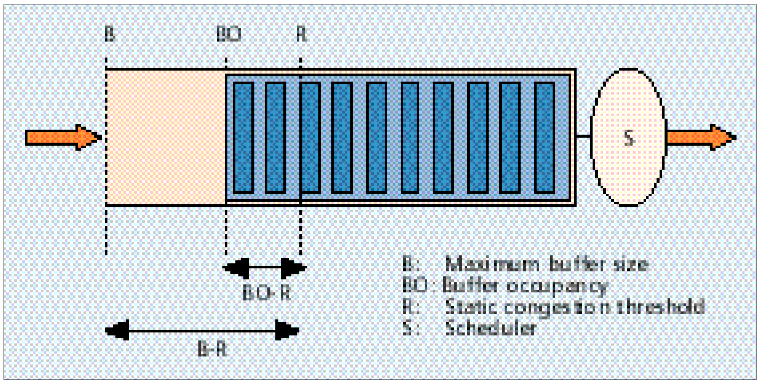


FIGURE 9. The FIFO-based with frame tagging and FBA (per-VC queuing).

$$BO > R \text{ and } \left[Q_i \left(\frac{N_a}{BO} \right) > Z \left(\frac{B-R}{BO-R} \right) \right]$$

where R is a static congestion threshold, Q_i is the number of cells that reside in the buffer of the ith connection, N_a is the number of active connections, Z is a scaling factor, and B is the maximum buffer size. The threshold R determines when the buffer is congested. If the BO is less than the threshold R, then no frames are discarded. When the BO is greater than R an incoming frame is discarded if the connection has more than its fair share of the buffer space. The discarding decision is determined in the second condition, where

$$\frac{BO}{N_a}$$

is a measure of the fair allocation of each connection. Furthermore, it should be noted that

$$Z \frac{B-R}{BO-R}$$

decreases as the free buffer space is reduced. The more the buffer is congested, the more likely it is that the incoming frames will be discarded. Simulation results showed that with identical TCP sources there is a trade off between efficient use of the buffer and its fair allocation to the TCP sources [7]. Fig. 9 illustrates how the parameters of the above mechanism are used.

Basak and Pappu in [24] used a modified FBA scheme, initially proposed in [25], the weighted FBA, which allows weighted sharing of buffer space among VCs. Each time a new frame arrives, a congestion threshold, Thrss(t)[4], is assigned to the connection. The value of Thrss(t) is a function of the unused buffer space and it is used in the packet discard decision of the incoming flow. Two more thresholds, LBO and HBO, are derived as a fixed fraction of the Thrss(t). The simulations that were carried out showed that the FBA scheme can share the buffer space in accordance with the weights but that the sharing is not tight enough. Hence, it was suggested that this mechanism fails to share the buffer space in proportion to the connection MCRs.

FIFO-BASED WITH FRAME TAGGING AND WBA (PER-VC QUEUING)

Goyal et al. in [26] suggested the use of dynamic thresholds, i.e., thresholds that are adjusted according to the buffer occupancy. The buffer management scheme that was proposed is called weighted buffer allocation (WBA). The buffer is divided into two parts, loaded and unloaded, with the use of a threshold R. Per-VC queuing is also implemented and thresholds are set in accordance with the connection MCR. When the buffer is unloaded, all the frames are accepted. In case of congestion, the frames may be dropped. Tagged cells are stored only if the total number of cells in the buffer is less than a dynamic threshold. The discarding decision is based on the following formula:

$$(Q_i - R \cdot W_i) \cdot N_a < Z \cdot (BO - R)$$

where Q_i is the number of cells that reside in the buffer of the ith connection, W_i is a per-VC weight set in accordance with the connection MCR, N_a is the number of active connections, Z is a scaling factor, and BO is the buffer occupancy. Furthermore, the decision to store tagged cells in the buffer is also based on a dynamic threshold defined from:

$$L_i < R \cdot W_i$$

where L_i is the number of tagged cells residing in the buffer. Obviously, the more the buffer occupancy exceeds the threshold R, the more difficult it is for a tagged cell to be stored. In this way, the number of tagged cells is controlled.

To summarize, only if there is enough space in the buffer and there are few tagged cells residing in the buffer can more tagged cells be stored. Furthermore, other drop policies such as EPD can be used in conjunction with WBA when severe congestion occurs.

Goyal et al. concluded that their mechanism is sufficient to provide GFR guarantees only if the sources transmit equal amounts of traffic.

FIFO-BASED WITH DFBA

Goyal et al. in [27] investigated whether FIFO-based mechanisms can provide GFR service guarantees to SACK TCP traffic in networks with low link utilization. Multiple TCP connections were multiplexed on each VC but the cells of different frames within the connection were not interleaved. A complex buffer management mechanism, called differential fair buffer allocation (DFBA), was implemented for controlling the per-VC buffer occupancy. The DFBA scheme drops cells in a probabilistic way according to static per-VC weights, thresholds, and number of cells within the buffer. There is no frame policy implementation, i.e., all the cells are untagged. The simulations showed that the TCP throughput could be proportional to the buffer allocation for each connection indicating that GFR guarantees could be provided.

Goyal et al. also showed that the TCP throughput, which is dependent on the TCP window size and the round trip time (RTT), could be artificially controlled by dropping frames at specific values of the window size. However, in practice the network elements cannot control the TCP window size because they are unaware of its dynamically changing value. Nevertheless, the network elements can estimate the TCP throughput by counting the number of cells stored in the output buffer.

FIFO-BASED WITH VIRTUAL QUEUING

Siu et al. proposed an EPD variant using the virtual queuing (VQ) technique [28]. Virtual queuing emulates the round-robin (RR) buffer allocation provided by per-VC queuing on a shared FIFO queue. According to this technique, instead of

---

[4] The congestion threshold, Thrss(t), is a function of time t as it is evaluated each time a new frame arrives at the ATM switch buffer.

maintaining one queue per VC in the ATM switch buffer, a single FIFO queue is used where a separate "virtual" queue is implemented for each VC_i by maintaining a state variable M_i for each VC_i (per-VC accounting). Each M_i is incremented by one when a cell arrives from VC_i at the FIFO queue, but is decreased in a round-robin fashion when a cell departs from VC_i, as if per-VC queuing and round-robin scheduling disciplines had been implemented. The EPD mechanism is then applied to each "virtual" queue in the same manner as for a normal queue.

The simulations conducted by Siu et al. used both TCP and UDP sources and concluded that minimum cell rate guarantees could be supported by the above mechanism, also allowing for the fair access of the remaining bandwidth among the sources.

RR AND WRR SCHEDULING DISCIPLINES WITH PER-VC QUEUING

The simplest (per-VC) scheduling discipline is the round robin (RR), according to which all the queues are served in a fixed order. If there are cells stored in a served queue then the outgoing link will be used by the connection up to a predetermined time duration. If the queue is empty, the scheduler instantly serves another queue. Traffic parameters, such as MCR, are not taken into account. Therefore, each VC occupies a fraction of the outgoing link bandwidth, which depends on the overall traffic of the VC trying to transmit.

Intuitively, the RR scheduler cannot support the GFR service, where each connection should be guaranteed at least the MCR rate. Pappu and Basak in [15] showed with simulations that RR is unable to support GFR guarantees. In the same contribution, however, a more complex scheduler, called weighted round robin (WRR), was simulated. With WRR, each connection has a weight that is set proportionally to the respective MCR, and is served accordingly by the scheduler. If all weights are equal, WRR is equivalent to RR, as depicted in Fig. 10. When the queue length is below the LBO, the scheduler divides the entire bandwidth according to the ratio of the weights of the VCs. If the queue length exceeds the LBO threshold, but is still below the HBO, the MCR is guaranteed to each connection since only the untagged cells corresponding to each connection's MCR are admitted into the buffer. The simulations showed that WRR can provide GFR service guarantees.

Huang et al. in [29] also investigated the performance of a GFR implementation using WRR scheduling. The buffer was divided in a similar manner as in the simulation study mentioned previously. However, per-VC and per-port thresholds[5] were considered. It was shown that per-VC buffer allocation and WRR scheduling are sufficient to support rate guarantees provided that the buffer is not small. On the contrary, when per-port buffer allocation was used, rate guarantees were not provided, especially with the simultaneous presence of a greedy source.[6] The use of network-based frame tagging did not affect the simulation results.

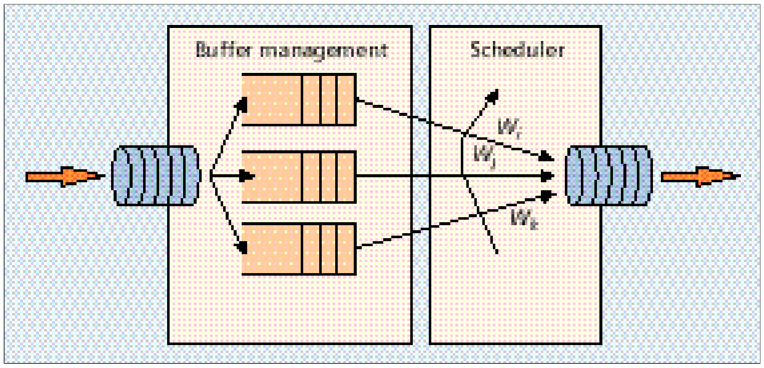


FIGURE 10. The Weighted Round-Robin (WRR) scheduling discipline.

WRR SCHEDULING DISCIPLINE WITH FBA AND WITHOUT FRAME TAGGING

Basak and Pappu in [24] also examined whether tagging could be substituted with a buffer allocation scheme. They simulated a mechanism that combines WRR and FBA. The role of the FBA was to prevent high-rate sources from blocking other sources using the buffer. Thus, the FBA scheme guaranteed that the buffer space was divided in proportion to the MCR. The simulations showed that rate guarantees could be supported.

WEIGHTED FAIR QUEUING-BASED SCHEDULING

Weighted fair queuing (WFQ), like WRR, is an approximation of the generalized processor sharing (GPS) scheduling [30]. WFQ can be generally used to give performance guarantees to connections carrying best-effort packet traffic, where each connection can be guaranteed bandwidth in proportion to its weight and in a fair manner. A number of variants of WFQ-based implementations exist, their main difference being their complexity.

Similarly to WRR, WFQ is capable of providing GFR guarantees, provided it is coupled with the proper buffer management mechanism. Bonaventure in [14] simulated the WFQ-based scheduling discipline found in [4]. He showed that the combination of WFQ-like per-VC scheduling and per-VC accounting could improve TCP throughput performance. Provided there was no cell loss, each source achieved full use of its allocated bandwidth. In all other cases, the TCP traffic sources face difficulties in using their "reserved" bandwidth.

Additional related papers studying the GFR service include [31-33].

ALTERNATIVE SOLUTIONS

Apart from GFR, the ABR, with MCR > 0, and the VBR, with service conformance definitions VBR.2 and VBR.3, service categories are also able to achieve minimum rate guarantees.

The ABR service category implements a rate-based feedback congestion-control mechanism to provide a loss-free best-effort service enhanced with a minimum guaranteed bandwidth [3]. ABR requires that all source and destination end systems attached to the ATM network implement the ABR congestion control mechanism [34] and adjust their transmission rates in response to feedback received from the network. Feedback information is conveyed by resource management (RM) cells that are transmitted by the source end systems, processed, and possibly modified by the intermediate ATM switches to indicate their resource availability and con-

---

[5] In a per-port buffer allocation scheme, the used (per-port) thresholds are common for all the connections sharing the same ATM port. In a per-VC buffer allocation scheme different thresholds are assigned to each connection.

[6] A source is called greedy when it has no flow control, i.e., a non-reactive source.

gestion state, and received by the destination end system, which in turn transmits them back to the source end system. ABR's advantage over GFR is its ability to identify the rate available to a connection on its bottleneck link, thus avoiding unnecessary consumption of network resources by allocating more than this amount to a connection on other links. On the other hand, GFR relies on the end-to-end flow control of the transport protocols such as TCP to control the consumption of the available network resources.

The authors believe that although simulation studies have found ABR to be as efficient as GFR [19], ABR's complexity, and the need for a substantial upgrade to the already installed base of ATM switches to support it, make it a less preferred choice than GFR at the moment.

VBR, with VBR.2 (tagging not enabled) and VBR.3 (tagging enabled), can also provide minimum cell rate guarantees to CLP = 0 traffic where the sustainable cell rate (SCR) parameter corresponds to the guaranteed minimum rate. A number of ATM Forum contributions [35, 36] proposed that GFR would become a simple extension of the VBR service category, via a new conformance definition, namely VBR.4 and the incorporation of suitable mechanisms. Simulation-based comparison studies between GFR and VBR.3 showed that GFR is generally superior to VBR.3 [12]. Finally, it is worth mentioning that the IETF recommends the nrt-VBR and ABR as the methods for supporting large aggregated traffic volumes between routers. This is likely to lead to implementations of a GFR-like nrt-VBR service within the context of UNI 4.0 signalling [37].

CONCLUSIONS

GFR is a major enhancement to the UBR service category that provides a guaranteed minimum cell rate to AAL5 frames for all existing connections and under all possible conditions. The ATM network treats identical connections, i.e., connections with the same QoS requirements and traffic contract, in the same way and provides similar final service, i.e., fairness.

We presented a number of representative mechanisms that have been proposed for providing GFR guarantees. The FIFO-based mechanisms were shown to be able to support the GFR service, although this depends greatly on the buffer management and policing mechanisms with which they are coupled. On the other hand, rate-guaranteeing scheduling disciplines, such as weighted round robin and weighted fair queuing, are capable of supporting the GFR service but require more complex implementations.

By using some of the above mechanisms, it is possible to provide GFR guarantees to VCs that need this type of service. However, most of the research concentrated on providing these guarantees, locally, per-switch. Further work is needed to verify whether end-to-end GFR guarantees can be achieved with the current implementations when there exist multiple ATM switches along the established path. Multiplexing several TCP flows onto a single VC, with high MCR and fairness in the allocation of the available bandwidth to the excess traffic, are other important topics to be examined. Verification of whether GFR is suitable for environments with different versions of TCP, for other transport protocols as well as for non-adaptive traffic, are also open issues.

GFR will benefit applications where big volumes of aggregate internetwork traffic are transferred, as in the case of LAN interconnection. Whether GFR will succeed against VBR and ABR in achieving what it has been designed for is still to be determined.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers of this article for their insightful comments and constructive criticism.

REFERENCES

[1] W. R. Stevens, *TCP/IP Illustrated, Vol. 1: The Protocols*, Addison-Wesley, 1994.
[2] J. Postel, Transmission Control Protocol, RFC793, Sept. 1981.
[3] ATM Forum, ATM Traffic Management Specification, Version 4.0, April 1996.
[4] R. Guerin and J. Heinanen, UBR+ Service Category Definition, ATM Forum contribution ATM96-1598, Dec. 1996.
[5] ATM Forum, Draft TM 4.1 Traffic Management Specification, March 1999.
[6] A. Romanow and S. Floyd, Dynamics of TCP Traffic over ATM Networks, *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 4, May 1995, pp. 633-641.
[7] R. Goyal *et al*., UBR+: Improving Performance of TCP over ATM-UBR Service, *Proceedings of ICC '97*, June 1997.
[8] R. Guerin and J. Heinanen, UBR+ Enhancements, ATM Forum contribution, ATM97-0015, Feb. 1997.
[9] R. Atkinson, Default IP MTU for use over ATM AAL5, RFC1626, May 1994.
[10] ATM Forum, ATM User Network Network Interface (UNI) Signalling Specification, Version 4.0, July 1996.
[11] B. Lee, GFR Dimensioning for TCP Traffic, ATM Forum contribution ATM98-0271, Feb. 1998.
[12] F. Hellstrand and A. Veres, Simulation of TCP/IP Router Traffic over ATM using GFR and VBR.3, ATM Forum contribution, ATM98-0087, Feb. 1998.
[13] A. Varma and D. Stiliadis, Hardware Implementation of Fair Queuing Algorithms for Asynchronous Transfer Mode Networks, *IEEE Commun. Mag*., Dec. 1997.
[14] O. Bonaventure, A simulation study of TCP with the Proposed GFR Service Category, DAGSTUHL Seminars 9725, High Performance Networks for Multimedia Applications, June 1997.
[15] S. Pappu and D. Basak, TCP over GFR Implementation with Different Service Disciplines: A Simulation Study, ATM Forum contribution ATM97-0310, April 1997.
[16] V. Jacobson, Congestion Avoidance and Control, *Proceedings of SIGCOMM '98*, Aug. 1988, pp. 314-329.
[17] W. R. Stevens, TCP Slow Start, Congestion Avoidance, Fast Retransmit and Fast Recovery Algorithms, Internet RFC2001, Jan. 1997.
[18] K. Fall and S. Floyd, Simulation-Based Comparisons of Tahoe, Reno, and SACK TCP, *Computer Communication Review*, vol. 26, no. 3, July 1996, pp. 5-21.
[19] O. Bonaventure, Providing Bandwidth Guarantees to Internetwork Traffic in ATM Networks, *Proceedings of IEEE ATM98 Workshop*, Fairfax, USA, May 1998.
[20] M. Mathis *et al*., TCP Selective Acknowledgement Options, RFC2018, Oct. 1996.
[21] R. Goyal *et al*., Providing Rate Guarantees to TCP over the ATM GFR Service, *Proceedings of 23rd Annual Conference on Local Computer Networks 1998* (LCN '98), Lowell, MA, Oct. 1998, pp. 390-398.
[22] R. Guerin *et al*., Scalable QoS Provision through Buffer Management, Sigcomm '98.
[23] J. Heinanen and K. Kilkki, A Fair Buffer Allocation Scheme, *Computer Communications*, Elsevier Science, vol.21, no.3, March 1998, pp. 220-226.
[24] D. Basak and S. Pappu, GFR Implementation Alternatives with Fair Buffer Allocation Schemes, ATM Forum contribution ATM97-0528, July 1997.
[25] A. K. Choudhary and E. L. Hahne, Dynamic Queue Length Thresholds in a Shared Memory ATM Switch, IEEE INFOCOM, March 1996, pp. 679-687.
[26] R. Goyal *et al*., Simulation Experiments with Guaranteed Frame Rate for TCP/IP Traffic, ATM Forum contribution ATM97-0607, July 1997.

[27] R. Goyal *et al.*, GFR: Providing Rate Guarantees with FIFO Buffers to TCP traffic, ATM Forum contribution ATM97-0831, Sept. 1997.
[28] K. Y. Siu, Y. Wu, and W. Ren, Virtual Queuing Techniques for UBR+ Service in ATM with Fair Access and Minimum Bandwidth Guarantee, *Proceedings of GLOBECOM '97*, 1997, pp. 1081-1085.
[29] J. Huang, B-J. Lee, and S. Khorsandi, A Simulation Study of GFR Implementation, ATM Forum contribution, ATM97-1035, Dec. 1997.
[30] A. Demers, S. Keshav, and S. Shenker, Design and Analysis of a Fair Queuing Algorithm, *Proceedings of SIGCOMM '89*, Austin, Sept. 1989.
[31] O. Eloumi and H. Afifi, Evaluation of FIFO-based Buffer Management Algorithms for TCP over Guaranteed Frame Rate Service, *Proceedings of ATM '98*, May 1998.
[32] M. Nabeshima, N. Yamanaka, and H. Hasegawa, New Scheduling Mechanisms which Efficiently Utilize Policing for GFR Service, *Proceedings of ATM '98*, May 1998.
[33] D. Wu and J. Chao, TCP/IP over ATM-GFR, *Proceedings of ATM '98*, May 1998.
[34] F. Bonomi and K. Fendick, The Rate-Based Flow Control Framework for the Available Bit Rate ATM Service, *IEEE Network*, vol. 9, no. 4, March/April 1995.
[35] J. Kenney, Satisfying UBR+ Requirements via a New VBR Conformance Definition, ATM Forum contribution, ATM97-0185, Feb. 1997.
[36] K. Fendick *et al.*, GFR as VBR Service, ATM Forum contribution, ATM98-0160, Feb. 1998.
[37] M. Maher, ATM Signalling Support for IP over ATM-UNI Signalling 4.0 Update, RFC2331, April 1998.

BIOGRAPHIES

ILIAS ANDRIKOPOULOS (I.Andrikopoulos@ee.surrey.ac.uk) received a diploma in physics from the University of Athens in 1993, and an M.Sc. in information technology from University College London in 1994. He joined the University of Surrey in 1995, where he is currently acting as a research fellow in the networks research group at the Centre for Communication Systems Research, while pursuing his PhD in the area of QoS provisioning in IP networks. His interests include TCP/IP over ATM, Internet technologies for multimedia applications, TCP congestion control, traffic management, scheduling and buffer management for high-speed networks, and network management.

ATHANASSIOS LIAKOPOULOS (aliako@netmode.ece.ntua.gr) received a diploma in electrical engineering from the National Technical University of Athens in 1996 and an MSc in Telematics from the University of Surrey in 1998. He worked for approximately two years at the Network Management Centre of NTUA, working in the areas of TCP/IP-based networks, information services, and protocols. His main interests include broadband telecommunication systems, particularly ATM networks and telecommunication network management.

GEORGE PAVLOU (G.Pavlou@ee.surrey.ac.uk) received a diploma in electrical engineering from the National Technical University of Athens. He obtained MSc and PhD degrees in computer science from University College London, where he served as a senior research fellow and lecturer. He is currently professor of information networking at the Centre for Communication Systems Research at the University of Surrey, where he leads the activities of the networks research group. Routing, quality of service, and network management and control are among his key research areas of interest.

ZHILI SUN (Z.Sun@ee.surrey.ac.uk) received a degree in mathematics from Nanjing University, China, and M.Phil. and Ph.D. degrees in computing from the University of Lancaster. He worked as a research fellow in the Telecom Research Group at QMW, and then as a research fellow and a senior research fellow at the Centre for Satellite Engineering Research of the University of Surrey. He is currently a lecturer at the Centre for Communication Systems Research at the University of Surrey. His research interests include terrestrial broadband communications and networks, and ATM interworking and Internet technologies for multimedia applications.