

Network Working Group  
Request for Comments: 2381  
Category: Standards Track

M. Garrett  
Bellcore  
M. Borden  
Bay Networks  
August 1998

## Interoperation of Controlled-Load Service and Guaranteed Service with ATM

### Status of this Memo

This document specifies an Internet standards track protocol for the Internet community, and requests discussion and suggestions for improvements. Please refer to the current edition of the "Internet Official Protocol Standards" (STD 1) for the standardization state and status of this protocol. Distribution of this memo is unlimited.

### Copyright Notice

Copyright (C) The Internet Society (1998). All Rights Reserved.

### Abstract

This document provides guidelines for mapping service classes, and traffic management features and parameters between Internet and ATM technologies. The service mappings are useful for providing effective interoperation and end-to-end Quality of Service for IP Integrated Services networks containing ATM subnetworks.

The discussion and specifications given here support the IP integrated services protocols for Guaranteed Service (GS), Controlled-Load Service (CLS) and the ATM Forum UNI specification, versions 3.0, 3.1 and 4.0. Some discussion of IP best effort service over ATM is also included.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [1]. (Note, in many cases the use of "MUST" or "REQUIRED" reflects our interpretation of the requirements of a related standard, e.g., ATM Forum UNI 4.0, rsvp, etc.)

## Table of Contents

|       |  |    |
|-------|--|----|
| 1.0   | Introduction .....   | 3  |
| 1.1   | General System Architecture .....                                    | 4  |
| 1.2   | Related Documents .....  | 7  |
| 2.0   | Major Protocol Features for Traffic Management and QoS .....         | 8  |
| 2.1   | Service Category and Bearer Capability .....                         | 8  |
| 2.1.1 | Service Categories for Guaranteed Service .....                      | 9  |
| 2.1.2 | Service Categories for Controlled Load .....                         | 10 |
| 2.1.3 | Service Categories for Best Effort .....                             | 11 |
| 2.2   | Cell Loss Priority Bit, Tagging and Conformance Definitions ..       | 11 |
| 2.3   | ATM Adaptation Layer .....   | 13 |
| 2.4   | Broadband Low Layer Information .....                                | 13 |
| 2.5   | Traffic Descriptors .....  | 13 |
| 2.5.1 | Translating Traffic Descriptors for Guaranteed Service ..            | 15 |
| 2.5.2 | Translating Traffic Descriptors for Controlled Load<br>Service ..... | 18 |
| 2.5.3 | Translating Traffic Descriptors for Best Effort Service ..           | 19 |
| 2.6   | QoS Classes and Parameters .....                                     | 19 |
| 2.7   | Additional Parameters -- Frame Discard Mode .....                    | 22 |
| 3.0   | Additional IP-Integrated Services Protocol Features .....            | 22 |
| 3.1   | Path Characterization Parameters for IP Integrated Services ..       | 22 |
| 3.2   | Handling of Excess Traffic .....                                     | 24 |
| 3.3   | Use of Guaranteed Service Adspec Parameters and Slack Term ..        | 25 |
| 4.0   | Miscellaneous Items .....  | 26 |
| 4.1   | Units Conversion .....   | 26 |
| 5.0   | Summary of ATM VC Setup Parameters for Guaranteed Service .....      | 27 |
| 5.1   | Encoding GS Using Real-Time VBR .....                                | 28 |
| 5.2   | Encoding GS Using CBR .....  | 29 |
| 5.3   | Encoding GS Using Non-Real-Time VBR .....                            | 30 |
| 5.4   | Encoding GS Using ABR .....  | 30 |
| 5.5   | Encoding GS Using UBR .....  | 30 |
| 5.6   | Encoding GS Using UNI 3.0 and UNI 3.1. ....                          | 31 |
| 6.0   | Summary of ATM VC Setup Parameters for Controlled Load Service ..    | 32 |
| 6.1   | Encoding CLS Using Non-Real-Time VBR .....                           | 32 |
| 6.2   | Encoding CLS Using ABR .....   | 33 |
| 6.3   | Encoding CLS Using CBR .....   | 35 |
| 6.4   | Encoding CLS Using Real-Time VBR .....                               | 35 |
| 6.5   | Encoding CLS Using UBR .....   | 35 |
| 6.6   | Encoding CLS Using UNI 3.0 and UNI 3.1. ....                         | 35 |
| 7.0   | Summary of ATM VC Setup Parameters for Best Effort Service .....     | 36 |
| 7.1   | Encoding Best Effort Service Using UBR .....                         | 37 |
| 8.0   | Security Considerations .....  | 38 |
| 9.0   | Acknowledgements .....   | 38 |
|       | Appendix 1 Abbreviations .....                                       | 39 |
|       | References .....   | 40 |
|       | Authors' Addresses .....   | 42 |
|       | Full Copyright Statement .....                                       | 43 |

## 1.0 Introduction

We consider the problem of providing IP Integrated Services [2] with an ATM subnetwork. This document is intended to be consistent with the rsvp protocol [3] for IP-level resource reservation, although it applies also in the general case where GS and CLS services are supported through other mechanisms. In the ATM network, we consider ATM Forum UNI Signaling, versions 3.0, 3.1 and 4.0 [4, 5, 6]. The latter uses the more complete service model of the ATM Forum's TM 4.0 specification [7, 8].

This is a complex problem with many facets. In this document, we focus on the service types, parameters and signalling elements needed for service interoperation. The resulting service mappings can be used to provide effective end-to-end Quality of Service (QoS) for IP traffic that traverses ATM networks.

The IP services considered are Guaranteed Service (GS) [9] and Controlled Load Service (CLS) [10]. We also discuss the default Best Effort Service (BE) in parallel with these. Our recommendations for BE are intended to be consistent with RFC 1755 [11], and [12], which define how ATM VCs can be used in support of normal BE IP service. The ATM services we consider are:

|        |                                 |
|--------|---------------------------------|
| CBR    | Constant Bit Rate               |
| rtVBR  | Real-time Variable Bit Rate     |
| nrtVBR | Non-real-time Variable Bit Rate |
| UBR    | Unspecified Bit Rate            |
| ABR    | Available Bit Rate              |

In the case of UNI 3.x signalling, where these service are not all clearly distinguishable, we identify the appropriate available services.

We recommend the following service mappings, since they follow most naturally from the service definitions:

|                    |    |  |
|--------------------|----|--|
| Guaranteed Service | -> | CBR or rtVBR                             |
| Controlled Load    | -> | nrtVBR or ABR (with a minimum cell rate) |
| Best Effort        | -> | UBR or ABR                               |

For completeness, however, we provide detailed mappings for all service combinations in Sections 5, 6, 7 and identify how each meets or fails to meet the requirements of the higher level IP services. The reason for not restricting mappings to the most obvious or natural ones is that we cannot predict how widely available all of these services will be as ATM deployment evolves. A number of

differences in service definitions, such as the treatment of packets in excess of the flow traffic descriptor, make service mapping a relatively complicated subject.

The remainder of this introduction provides a general discussion of the system configuration and other assumptions. Section 2 considers the relevant ATM protocol elements and the corresponding features of Guaranteed, Controlled Load and Best Effort services (the latter being the default "service"). Section 3 discusses a number of remaining features of the IP services and how they can be handled on an ATM subnetwork. Section 4 addresses the conversion of traffic descriptors to account for ATM-layer overheads. Section 5 gives the detailed VC setup parameters for Guaranteed Service, and considers the effect of using each of the ATM service categories. Section 6 provides a similar treatment for Controlled Load Service. Section 7 considers Best Effort service.

This document is only a part of the total solution to providing the interworking of IP integrated services with ATM subnetworks. The important issue of VC management, including flow aggregation, is considered in [13, 14, 15]. We do not consider how routing, QoS sensitive or not, interacts with the use of ATM VCs. We expect that a considerable degree of implementation latitude will exist, even within the guidelines presented here. Many aspects of interworking between IP and ATM will depend on economic factors, and will not be subject to standardization.

## 1.1 General System Architecture

We assume that the reader has a general working knowledge of IP, RSVP and ATM protocols. The network architecture we consider is illustrated in Figure 1. An IP-attached host may send unicast datagrams to another host, or may use an IP multicast address to send packets to all of the hosts which have "joined" the multicast "tree". In either case, a destination host may then use RSVP to establish resource reservation in routers along the internet path for the data flow.

An ATM network lies in the path (chosen by the IP routing), and consists of one or more ATM switches. It uses SVCs to provide both resources and QoS within the ATM cloud. These connections are set up, added to (in the case of multipoint trees), torn down, and controlled by the edge devices, which act as both IP routers and ATM interfaces, capable of initiating and managing VCs across the ATM user-to-network (UNI) interface. The edge devices are assumed to be fully functional in both the IP int-serv/RSVP protocols and the ATM UNI protocols, as well as translating between them.

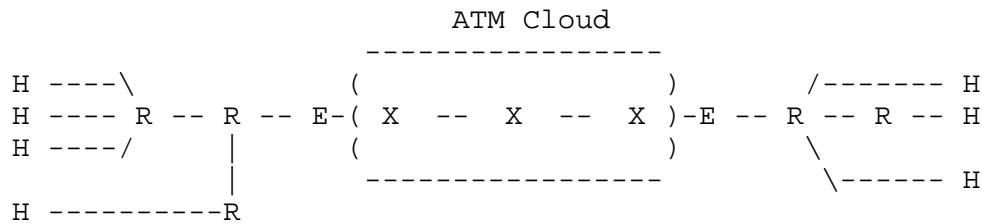


Figure 1: Network Architecture with Hosts (H), Routers (R), Edge Devices (E) and ATM Switches (X).

When considering the edge devices with respect to traffic flowing from source to destination, the upstream edge device is called the "ingress" point and the downstream device the "egress" point. The edge devices may be considered part of the IP internet or part of the ATM cloud, or both. They process RSVP messages, reserve resources, and maintain soft state (in the control path), and classify and schedule packets (in the data path). They also initiate ATM connections by signalling, and accept or refuse connections signalled to them. They police and schedule cells going into the ATM cloud. The service mapping function occurs when an IP-level reservation (RESV message) triggers the edge device to translate the RSVP service requirements into ATM VC (UNI) semantics.

A range of VC management policies are possible, which determine whether a flow initiates a new VC or joins an existing one. VCs are managed according to a combination of standards and local policy rules, which are specific to either the implementation (equipment) or the operator (network service provider). Point-to-multipoint connections within the ATM cloud can be used to support general IP multicast flows. In ATM, a point to multipoint connection can be controlled by the source (or root) node, or a leaf initiated join (LIJ) feature in ATM may be used. The topic of VC management is considered at length in other ISSLL documents [13, 14, 15].

Figure 2 shows the functions of an edge device, summarizing the work not part of IP or ATM abstractly as an InterWorking Function (IWF), and segregating the control and data planes.

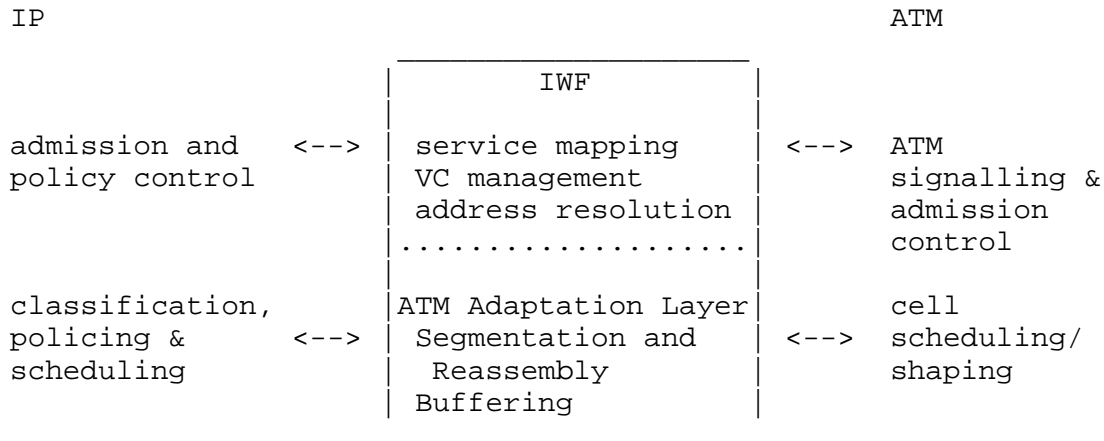


Figure 2: Edge Device Functions showing the IWF

In the logical view of Figure 2, some functions, such as scheduling, are shown separately, since these functions are present on both the IP and ATM sides. However it may be possible in an integrated implementation to combine such functions.

The service mapping and VC management functions can be highly interdependent. For example: (i) Multiple integrated-services flows may be aggregated to use one point-to-multipoint VC. In this case, we assume the IP flows are of the same service type and their parameters have been merged appropriately. (ii) The VC management function may choose to allocate extra resources in anticipation of further reservations or based on an empiric of changing TSpecs. (iii) There MUST exist a path for best effort flows and for sending the rsvp control messages. How this interacts with the establishment of VCs for QoS traffic may alter the desired characteristics of those VCs. See [13, 14, 15] for further details on VC management.

Therefore, in discussing the service mapping problem, we will assume that the VC management function of the IWF can always express its result in terms of an IP-level service with some QoS and TSpec. The service mapping algorithm can then identify the appropriate VC parameters as if a new VC were to be created for this service. The VC management function can then use this information consistent with its own policy, which determines whether the resulting action uses new or existing VCs.

## 1.2 Related Documents

Earlier ATM Forum documents combined signalling, traffic management and other areas into a single document, e.g., UNI 3.0 [4] and UNI 3.1 [5]. The 3.1 release was used to correct errors and fix alignment with the ITU. While UNI 3.0 and 3.1 are incompatible in terms of actual codepoints, the semantics are generally the same. Therefore, we will often refer to UNI 3.x to mean either version of the ATM protocol.

After 3.1, the ATM Forum released documents separately for each technical working group. The UNI Signalling 4.0 [6] and Traffic Management 4.0 [7] documents represent a consistent overall ATM protocol, and we will sometime refer to the combination as TM/UNI 4.0.

Within the IETF, related material includes the work of the rsvp [3], int-serv [2, 9, 10, 16, 17] and ion working groups [11, 12]. Rsvp defines the resource reservation protocol (which is analogous to signalling in ATM). Int-serv defines the behavior and semantics of particular services (analogous to the Traffic Management working group in the ATM Forum). Ion defines interworking of IP and ATM for traditional Best Effort service, and generally covers issues related to IP/ATM routing and addressing.

A large number of ATM signalling details are covered in RFC 1755 [10]; e.g., differences between UNI 3.0 and UNI 3.1, encapsulation, frame-relay interworking, etc. These considerations extend to IP over ATM with QoS as well. The description given in this document of IP Best Effort service (i.e. the default behavior) over ATM is intended to be consistent with RFC 1755 and its extension for UNI 4.0 [11], and those documents are to be considered definitive. For non-best-effort services, certain IP/ATM features will diverge from the following RFC 1755. We have attempted to note such differences explicitly. (For example, best effort VCs may be taken down on timeout by either edge device, while QoS VCs are only removed by the upstream edge device when the corresponding rsvp reservation is deleted.)

Another related document is RFC 1821 [17], which represents an early discussion of issues involved with interoperating IP and ATM protocols for integrated services and QoS.

## 2.0 Major Protocol Features for Traffic Management and QoS

The ATM Call Setup is sent by the ingress edge device to the ATM network to establish end-to-end (ATM) service. This setup contains the following information.

- Service Category/Broadband Bearer Capability
- AAL Parameters
- Broadband Low Layer Information
- Calling and Called Party Addressing Information
- Traffic Descriptors
- QoS Class and/or Parameters
- Additional Parameters of TM/UNI 4.0

In this section, we discuss each of these items as they relate to creating ATM VCs suitable for GS, CLS and BE services. We do not discuss routing and addressing at all, since they are (at least presently) independent of QoS. Following the section on service categories, we discuss tagging and conformance definitions for IP and ATM. These do not appear explicitly as set-up parameters in the above list, since they are implied by the policing method used.

### 2.1 Service Category and Bearer Capability

The highest level of abstraction distinguishing features of ATM VCs is in the service category or bearer capability. Service categories were introduced in TM/UNI 4.0; previously the bearer capability was used to discriminate at this level.

These elements indicate the general properties of a VC: whether there is a real-time delay constraint, whether the traffic is constant or variable rate, the applicable traffic and QoS description parameters and (implicitly) the complexity of some supporting switch mechanisms (e.g., ABR).

For UNI 3.0 and UNI 3.1, there are only two distinct options for bearer capabilities (in our context):

- BCOB-A: constant rate, timing required, unicast/multipoint;
- BCOB-C: variable rate, timing not required, unicast/multipoint.

A third capability, BCOB-X, can be used as a substitute for the above two capabilities, with its dependent parameters (traffic type and timing requirement) set appropriately. The distinction between the BCOB-X formulation and the "equivalent" (for our purposes) BCOB-A and BCOB-C constructs is whether the ATM network is to provide pure cell relay service or interwork with other technologies (with interoperable signalling), such as narrowband ISDN. Where this



distinction is applicable, the appropriate code SHOULD be used (see [5] and related ITU specs, e.g., I.371).

In TM/UNI 4.0 the service categories are:

- Constant Bit Rate (CBR)
- Real-time Variable Bit Rate (rtVBR)
- Non-real-time Variable Bit Rate (nrtVBR)
- Unspecified Bit Rate (UBR)
- Available Bit Rate (ABR)

The first two of these are real-time services, so that rtVBR is new to TM/UNI 4.0. The ABR service is also new to TM/UNI 4.0. UBR exists in all specifications, although it is called "best effort" in UNI 3.x. In either case it is indicated by the "best effort" indication flag (and the QoS Class set to 0).

The Service Category in TM/UNI 4.0 is encoded into the same signalled Information Element (IE) as the Bearer Capability in UNI 3.x, for the purpose of backward compatibility. Thus, we use the convention of referring to Service Category (CBR, rtVBR, nrtVBR, UBR, ABR) for TM/UNI 4.0 (where the bearer capability is implicit). When we refer to the Bearer Capability explicitly (BCOB-A, BCOB-C, BCOB-X), we are describing a UNI 3.x signalling message.

In principle, it is possible to support any service through the use of BCOB-A/CBR. This is because the CBR service is equivalent to having a "pipe" of a specified bandwidth. However, it may be significantly more efficient to use the other ATM services where applicable and available [17].

#### 2.1.1 Service Categories for Guaranteed Service

There are two possible mappings for GS:

- CBR (BCOB-A)
- rtVBR

Real-time support is REQUIRED for GS. Thus in UNI 3.x, the Bearer Class BCOB-A (or an equivalent BCOB-X formulation) MUST be used. In TM/UNI 4.0 either CBR or rtVBR is appropriate. The use of rtVBR may encourage recovery of allocated bandwidth left unused by a source. It also accommodates more bursty sources with a larger token bucket burst parameter, and permits the use of tagging for excess traffic (see Section 2.2).

Neither the BCOB-C Bearer Class, nor nrtVBR, UBR, ABR are good matches for the GS service. These provide no delay estimates and cannot guarantee consistently low delay for every packet.

For BCOB-A or CBR, specification of a peak cell rate (PCR) is REQUIRED by ATM standards. In these cases, PCR is the nominal clearing rate with a nominal jitter toleration (bucket size), CDVT.

When rtVBR is specified, two rates, PCR and SCR are REQUIRED (by ATM standards). This models bursty traffic with specified peak and sustainable rates. The corresponding ATM token bucket depth values are CDVT, and CDVT+BT, respectively.

### 2.1.2 Service Categories for Controlled Load

There are three possible good mappings for CLS:

- CBR (BCOB-A)
- nrtVBR (BCOB-C)
- ABR

Note that under UNI 3.x, there are equivalent services to CBR and nrtVBR, but not ABR. The first, with a CBR/BCOB-A connection, provides a higher level of QoS than is necessary, but it may be convenient to simply allocate a fixed-rate "pipe", which we expect to be ubiquitously supported in ATM networks. However unless this is the only choice available, it would probably be wasteful of network resources.

The nrtVBR/BCOB-C category is perhaps the best match, since it provides for allocation of bandwidth and buffers with an additional peak rate indication, similar to the CLS TSpec. Excess traffic can be handled by CLP bit tagging with VBR.

The ABR category with a positive MCR aligns with the CLS idea of "best effort with a floor." The ATM network agrees to forward cells with a rate of at least MCR, which MUST be directly converted from the token bucket rate of the receiver TSpec. The bucket size parameter measures approximately the amount of buffer necessary at the IWF. This buffer serves to absorb the bursts allowed by the token bucket, since they cannot be passed directly into an ABR VC.

The rtVBR category can be used, although the edge device MUST then determine values for CTD and CDV. Since there are no corresponding IP-level parameters, their values are set as a matter of local policy.

The UBR category does not provide enough capability for Controlled Load. The point of CLS is to allow an allocation of resources. This is facilitated by the token bucket traffic descriptor, which is unavailable with UBR.

### 2.1.3 Service Categories for Best Effort

All of the service categories have the capability to carry Best Effort service, but the natural service category is UBR (or, in UNI 3.x, BCOB-C or BCOB-X, with the best effort indication set). CBR or rtVBR clearly could be used, and since the service is not real-time, a nrtVBR connection could also be used. In these cases the rate parameter used reflects a bandwidth allocation in support of the ingress edge device's best effort connectivity to the egress edge router. It would be normal for traffic from many source/destination pairs to be aggregated on this connection; indeed, since Best Effort is the default IP behavior, the individual flows are not normally identified or accounted for. CBR may be a preferred solution in the case where best effort traffic is sufficiently highly aggregated that a simple fixed-rate pipe is efficient. Both CBR and nrt-VBR provide explicit bandwidth allocation which may be useful for billing purposes. In the case of UBR, the network operator SHOULD allocate bandwidth for the overall service through the admission control function, although such allocation is not done explicitly per VC.

An ABR connection could similarly be used to support Best Effort traffic. Indeed, the support of data communications protocols such as TCP/IP is the explicit purpose for which ABR was designed. It is conceivable that a separate ABR connection would be made for each IP flow, although the normal case would probably have all IP Best Effort traffic with a common egress router sharing a single ABR connection.

The rt-VBR service category may be considered less suitable, simply because both the real-time delay constraint and the use of SCR/BT add unnecessary complexity.

See specifications from the IETF ion working group [10, 11] for related work on support of Best Effort service with ATM.

## 2.2 Cell Loss Priority Bit, Tagging and Conformance Definitions

Each ATM cell header carries a Cell Loss Priority (CLP) bit. Cells with CLP=1 are said to be "tagged" or "marked" and have lower priority. This tagging may be done by the source, to indicate relative priority within the VC, or by a switch, to indicate traffic in violation of policing parameters. Options involving the use of tagging are decided at call setup time.

A Conformance Definition is a rule that determines whether a cell is conforming to the traffic descriptor of the VC. The conformance definition is given in terms of a Generic Cell Rate Algorithm (GCRA), also known as a "leaky bucket" algorithm, for CBR and VBR services. The conformance definition also specifies rules for tagging traffic in excess of the {SCR, MBS} GCRA traffic descriptor. (Note, the term "compliance" in ATM is used to describe the behavior of a connection, as opposed to "conformance", which applies to a single cell.)

The network may tag cells that are non-conforming, rather than dropping them if the VC set-up requests tagging and the network supports the tagging option. When tagging is used and congestion occurs, a switch MUST attempt to discard tagged cells in preference to discarding CLP=0 cells. However, the mechanism for doing this is completely implementation specific. The behavior that best meets the requirements of IP Integrated Services is where tagged cells are treated as "best effort" in the sense that they are transported when bandwidth is available, queued when buffers are available, and dropped when resources are overcommitted. ATM standards, however, do not explicitly specify treatment of tagged traffic. Providers of GS and CLS service with ATM subnetworks SHOULD ascertain the actual behavior of ATM implementation with respect to tagged cells.

Since GS and CLS services REQUIRE excess traffic to be treated as best effort, the tagging option SHOULD always be chosen (if supported) in the VC setup as a means of "downgrading" the cells comprising non-conformant packets. The term "best effort" can be interpreted in two ways. The first is as a service class that, for example, may be implemented as a separate queue. The other sense is more generic, meaning that the network makes a best effort to transport the traffic. A reasonable interpretation of this is that a network with no contending traffic would transport the packet, while a very congested network would drop the packet. A mechanism that tags best effort packets with lower loss priority (such as with the ATM CLP bit) would drop some of these packets, but would not reorder the remaining ones with respect to the conforming portion of the flow. The "best effort" mechanism for excess traffic does not necessarily have to be the same as that for best effort "service", as long as it fits this generic sense of best effort.

There are three conformance definitions of VBR service (for both rtVBR and nrtVBR) to consider. In VBR, only the conformance definition VBR.3 supports tagging and applies the GCRA with rate PCR to the aggregate CLP=0+1 cells, and another GCRA with rate SCR to the CLP=0 cells. This conformance definition SHOULD always be used with a VBR service supporting IP integrated services. For UBR service, conformance definition UBR.2 supports the use of tagging, but a CLP=1 cell does not imply non-conformance; rather, it may be used by the

network to indicate congestion.

In TM/UNI 4.0 tagging is not a feature of the conformance definitions for the CBR or ABR service categories. (Since conformance definitions are generally network specific, some implementations CBR or ABR may, in fact, use tagging in some way.) Wherever an ATM network does support tagging, in the sense of transporting CLP=1 cells on a "best effort" basis, it is a useful and preferable mechanism for handling excess traffic.

It is always better for the IWF to tag cells when it can anticipate that the ATM network would do so. This is because the IWF knows the IP packet boundaries and can tag all of the cells corresponding to a packet. If left to the ATM layer UPC, the network would inevitably drop some of the cells of a packet while carrying others, which would then be dropped by the receiver. Therefore, the IWF, knowing the VC GCRA parameters, SHOULD always anticipate the cells which will be tagged by the ATM UPC and tag all of the cells uniformly across each affected packet. See Section 3.2 for further discussion of excess traffic.

### 2.3 ATM Adaptation Layer

The AAL type 5 encoding SHOULD be used, as specified in RFC 1483 and RFC 1755. For AAL-5, specification of the maximum SDU size in both the forward and reverse directions is REQUIRED. Both GS and CLS specify a maximum packet size, M, as part of the TSpec and this value SHOULD be used (corrected for AAL headers) as the maximum SDU in each direction for unicast connections, and for unidirectional point-to-multipoint connections. When multiple flows are aggregated into a single VC, the M parameters of the receiver TSpecs are merged according to rules given in the GS and CLS specs.

### 2.4 Broadband Low Layer Information

The B-LLI Information Element is transferred transparently by the ATM network between the edge devices and is used to specify the encapsulation method. Multiple B-LLI IEs may be sent as part of negotiation. The LLC/SNAP encapsulation [18] SHOULD be supported as the default, but "null" or "VC encapsulation" MAY also be allowed. Implementations SHOULD follow RFC 1577 [19] and RFC 1755 [10] for BLLI usage.

### 2.5 Traffic Descriptors

The ATM traffic descriptor always contains a peak cell rate (PCR) (for each direction). For VBR services it also contains a sustainable cell rate (SCR) and maximum burst size (MBS). The SCR

and MBS form a leaky bucket pair (rate, depth), while the bucket depth parameter for PCR is CDVT. Note that CDVT is not signalled explicitly, but is determined by the network operator, and can be viewed as a measure of the jitter imposed by the network.

Since CDVT is generally presumed to be small (equivalent to a few cells of token bucket depth), and cannot be set independently for each connection, it cannot be used to account for the burstiness permitted by b of the IP-layer TSpec. Additional buffering may be needed at the IWF to account for the depth of the token bucket.

The ATM Burst Tolerance (BT) is equivalent to MBS (see TM 4.0 [6] for the exact equation). They are both expressions of the bucket depth parameter associated with SCR. The units of BT are time while the units of MBS are cells. Since both SCR and MBS are signalled, they can be computed directly from the IP layer traffic description. The specific manner in which resources are allocated from the traffic description is implementation specific. Note that when translating the traffic parameters, the segmentation overhead and minimum policed unit need to be taken into account (see Section 4.1 below).

In ATM UNI Signalling 4.0 there are the notions of Alternative Traffic Descriptors and Minimal Traffic Descriptors. Alternative Traffic Descriptors enumerate other acceptable choices for traffic descriptors and are not considered here. Minimal Traffic Descriptors are used in "negotiation," which refers to the specific way in which an ATM connection is set up. To illustrate, roughly, taking PCR as an example: A minimal PCR and a requested PCR are signalled, the requested PCR being the usual item signalled, and the minimal PCR being the absolute minimum that the source edge device will accept. When both minimal and requested parameters are present, the intermediate switches along the path may reduce the requested PCR to a "comfortable" level. This choice is part of admission control, and is therefore implementation specific. If at any point the requested PCR falls below the minimal PCR then the call is cleared. Minimal Traffic Descriptors can be used to present an acceptable range for parameters and ensure a higher likelihood of call admission. In general, our discussion of connection parameters assumes the values resulting from successful connection setup.

The Best Effort indicator (used only with UBR) and Tagging indicators (see Section 2.2) are also part of the signalled information element (IE) containing the traffic descriptor. In the UNI 4.0 traffic descriptor IE there is an additional parameter, the Frame Discard indicator, which is discussed below in Section 2.7.

### 2.5.1 Translating Traffic Descriptors for Guaranteed Service

For Guaranteed Service the source TSpec contains peak rate, rate and bucket depth parameters,  $p_s$ ,  $r_s$ ,  $b_s$ . The receiver TSpec contains corresponding parameters  $p_r$ ,  $r_r$ ,  $b_r$ . The (receiver) RSpec also has a rate,  $R$ . The two different TSpec rates are intended to support receiver heterogeneity, in the sense that receivers can accept different rates representing different subsets of the sender's traffic. Whenever rates from different receivers differ, the values MUST always be merged appropriately before being mapping into ATM parameters.

Note that when the sender and receiver TSpec rates  $r_s$ ,  $r_r$  differ, there is no mechanism specified (in either RSVP or the int-serv specs) for indicating which subset of the traffic is to be transported. Implementation of this feature is therefore completely network specific. The policing and scheduling mechanisms may simply be parameterized with the (lower) receiver rate, resulting in the random loss of traffic sufficient to make up the difference in rates.

The receiver TSpec rate describes the traffic for which resources are to be reserved, and may be used for policing, while the RSpec rate (which cannot be smaller) is used (perhaps in an implementation specific way) to modify the allocated service bandwidth in order to reduce the delay.

When mapping Guaranteed Service onto a rtVBR VC, the ATM traffic descriptor parameters (PCR, SCR, MBS) can be set canonically as:

```
PCR =  $p_r$ 
SCR =  $R$ 
MBS =  $b_r$ .
```

There are a number of conditions that may lead to different choices. The following discussion is not intended to set hard requirements, but to provide some interpretation and guidance on the bounds of possible parameter mappings. The ingress edge device generally includes a buffer preceding the ATM network interface. This buffer can be used to absorb bursts that fall within the IP-level TSpec, but not within the ATM traffic descriptor. The minimal REQUIREMENT for guaranteed service is that the delay in this buffer MUST NOT exceed  $b/R$ , and the delays within the ATM network MUST be accurately accounted for in the values of Adspec parameters C and D advertised by the ingress router (see Section 3.3 below).

If either an edge device buffer of size  $b_r$  exists or the ATM maximum burst size (MBS) parameter is at least  $b_r$ , then the various rate parameters will generally exhibit the following relationship:

$$r\_r \leq SCR \leq R \leq PCR \leq APB \leq \text{line rate}$$

$$r\_r \leq p\_r \leq APB$$

APB refers to the General Characterization Parameter, AVAILABLE\_PATH\_BANDWIDTH, which is negotiated in the Adspec portion of the PATH message. APB reflects the narrowest bottleneck rate along the path, and so is always no larger than the local line rate. The receiver SHOULD choose a peak rate no greater than APB for the reservation to be accepted, although the source peak rate,  $p\_s$ , could be higher, as the source does not know the value of APB. There is no advantage to allocating any rate above APB of course, so it is an upper bound for all the other parameters.

We might normally expect to find  $R \leq p\_r$ , as would be necessary for the simple mapping of  $PCR = p\_r$ ,  $SCR = R$  given above. However, a receiver is free to choose  $R > p\_r$  to lower the GS delay [8]. In this case, PCR cannot be set below R, because a burst of size b arriving into the buffer MUST be cleared at rate R to keep the first component of GS delay down to  $b/R$ . So here we will have  $PCR = R$ . It may seem that  $PCR = p\_r$  would be sufficient to avoid buffer overflow, since data is generated at the source at a rate bounded by  $p\_r$ . However, setting  $PCR < R$ , can result in the delay bound advertised by C and D not being met. Also, traffic is always subject to jitter in the network, and the arrival rate at a network element can exceed  $p\_r$  for short periods of time.

In the case  $R \leq p\_r$ , we may still choose PCR such that  $R \leq PCR < p\_r$ . The edge device buffer is then necessary (and sufficient) to absorb the bursts (limited to size  $b\_r + C\_sum + R D\_sum$ ) which arrive faster than they depart. For example, it may be the case that the cost of the ATM VC depends on PCR, while the cost of the Internet service reservation is not strongly dependent on the IP-level peak rate. The user may then have an incentive to set  $p\_r$  to APB, while the operator of the IP/ATM edge router has an incentive to reduce PCR as much as possible. This may be a realistic concern, since the charging models of IP and ATM are historically different as far as usage sensitivity, and the value of  $p\_r$ , if set close to APB, could be many times the nominal GS allocated rate of R. Thus, we can set PCR to R, with a buffer of size  $b\_r + C\_sum + R D\_sum$ , with no loss of traffic, and no violation of the GS delay bound.

A more subtle, and perhaps controversial case is where we set SCR to a value below R. The major feature of the GS service is to allow a receiver to specify the allocated rate R to be larger than the rate  $r\_r$  sufficient to transport the traffic, in order to lower the queueing delay (roughly) from  $b/r + C\_TOT/r + D\_TOT$  to  $b/R + C\_TOT/R + D\_TOT$ . To effectively allocate bandwidth R to the flow, we set SCR



to match  $R$ . (Note it is unnecessary in any case to set  $SCR$  above  $R$ , so the relation,  $SCR \leq R$ , is still true.) It is possible to set  $SCR$  to a value as low as  $r_r$ , without violating the delay bounds or overflowing the edge device buffer. With  $PCR = R$ , a burst of size  $b$  will be buffered and sent into the ATM network at rate  $R$ , so the last byte suffers delay only  $b/R$ . Any further traffic will be limited to rate  $r_r$ , which is  $SCR$ , so with the arriving and departing rates matched, its delay will also be no more than  $b/R$ .

While this scenario meets the GS service requirements, the penalty for allocating  $SCR = r_r$  rather than  $R$  is that the delay in the ATM network will have a component of  $MBS/SCR$ , which will be  $b/r$  rather than  $b/R$ , contained in the  $D$  term advertised for the ATM sub-network (see further discussion in Section 3.3 below). It is also true that allocating  $r$  instead of  $R$  in a portion of the path is rather against the spirit of GS. As mentioned above, this mapping may however be useful in practice in the case where pricing in the ATM network leads to different incentives in the tradeoff between delay and bandwidth than those of the user who buys IP integrated services.

Another point of view on parameter mapping suggests that  $SCR$  may merely reflect the traffic description, hence  $SCR = r_r$ , while the service requirement is expressed in the QoS parameter as  $CDV = b/R$ . Thus the ATM network may internally allocate bandwidth  $R$ , but it is free to use other methods as well to achieve the delay constraint. Mechanisms such as statistical flow/connection aggregation may be implemented in the ATM network and hidden from the user (or parameter mapping module in the edge router) which sees only the interface implemented in the signalled parameters.

Note that this discussion considers an edge device buffer size of  $b_r$ . In practice, it may be necessary for the AAL/segmentation module to buffer  $M$  bytes in converting packets to cells. Also an additional amount of buffer equal to  $C_{sum} + R D_{sum}$  is generally necessary to absorb jitter imposed by the upstream network [8].

With ATM, it is possible to have little or no buffer in the edge router, because the ATM VC can be set to accept bursts at peak rate. This may be unusual, since the edge router normally has enough buffer to absorb bursts according to the TSpec token bucket parameters. We consider two cases. First, if  $PCR \geq p_r$ , then  $MBS$  can be set to  $b_r$  and no buffering is necessary to absorb non-excessive bursts. The extra buffering needed to absorb jitter can also be transferred to  $MBS$ . This effectively moves the buffering across the UNI into the ATM network.

For completeness, we consider an edge router with no burst-absorbing buffers and an MBS parameter of approximately zero. In this case it is sufficient to set the rate parameters to  $PCR = SCR = \max(R, p_r)$ . This amounts to peak-rate allocation of bandwidth, which will not usually be very cost effective. This case may be relevant where the IP routers and ATM switches differ substantially in their buffering designs. IP-level users may typically specify much larger burst parameters than can be handled in the ATM subnet. Peak-rate bandwidth allocation provides a means to work around this problem. It is also true that intermediate tradeoffs can be formulated, where the burst-absorbing buffer is less than  $b$  bytes, and  $SCR$  is set above  $R$  and below  $p_r$ . Note that jitter-absorbing buffers ( $C_{sum} + R D_{sum}$ ) can not be avoided, generally, by increasing ATM rates, unless  $SCR$  is set to exceed the physical line rate(s) into the edge device for the flow.

For GS over CBR, the value of  $PCR$  may be mapped to the RSpec rate  $R$ , if the edge device has a buffer of size  $b_r + C_{sum} + R D_{sum}$ . With little or no burst buffering, the requirements resemble the zero-buffer case above, and we have  $PCR = \max(R, p_r)$ . Additional buffers sufficient to absorb network jitter, given by  $C_{sum}$ ,  $D_{sum}$ , MUST always be provided in the edge router, or in the ATM network via MBS.

### 2.5.2 Translating Traffic Descriptors for Controlled Load Service

The Controlled Load service TSpec has a peak rate,  $p$ , a "token bucket" rate,  $r$ , and a corresponding token bucket depth parameter,  $b$ . The receiver TSpec values are used to determine resource allocation, and a simple mapping for the nrtVBR service category is given by,

```
PCR = p_r
SCR = r_r
MBS = b_r.
```

The discussions in the preceding section on using edge device buffers to reduce  $PCR$  and/or  $MBS$  apply generally to the CLS over nrtVBR case as well. Extra buffers to accommodate jitter accumulated (beyond the  $b_r$  burst size allowed at the source) MUST be provided. For CLS, there are no Adspec parameters  $C$  and  $D$ , so the dimensioning of such buffers is an implementation design issue.

For ABR VCs, the TSpec rate  $r_r$  is used to set the minimum cell rate (MCR) parameter. Since there is no corresponding signalled bucket depth parameter, the edge device SHOULD have a buffer of at least  $b_r$  bytes, plus additional buffers to absorb jitter. With ABR, the ATM network can quickly throttle the actual transfer rate down to MCR, so a source transmitting above that rate can experience high loss at the

ingress edge device when the ATM network becomes congested.

For CBR, the TSpec rate  $r_r$  sets a lower bound on PCR, and again, the available buffering in the edge device SHOULD be adequate to accommodate possible bursts of  $b_r$ .

The REQUIREMENT for CLS that network delays approximate "best-effort service under unloaded conditions", is interpreted here to mean that it would be sufficient to allocate bandwidth resources so that the last byte of a burst of size  $b_r$  sees a delay approximately  $b_r/r_r$ . For example, a network element with no cross-traffic, a work conserving scheduler and an output link rate of  $r_L$ , might provide delays in the range from  $M/r_L$  to  $b_r/r_L$ , that are much lower than  $b_r/r_r$ . While the access to the full link bandwidth ( $r_L$ ), as reflected in this example, is a more literal interpretation of delay "under unloaded conditions" for a shared link, an ATM VC may only have access to bandwidth equal to its nominal allocation (some implementation specific function of SCR and PCR).

### 2.5.3 Translating Traffic Descriptors for Best Effort Service

For Best Effort service, there is no traffic description. The UBR service category allows negotiation of PCR simply to allow the source to discover the smallest physical bottleneck along the path. The ingress edge router may set PCR to the ATM line rate, and then when the VC setup is complete, the returned value indicates an upper bound on throughput. For UBR service, resources may be allocated for the overall service (i.e., not per-VC) using the (implementation specific) admission control features of the ATM switches.

Often a service provider will statically configure large VCs with a certain bandwidth allocation to handle all best effort traffic between two edge routers. ABR, CBR or nrtVBR VCs are appropriate for this design, with traffic parameters set to comfortably accommodate the expected traffic load. See IETF ION specifications for IP over ATM signalling [10, 11].

### 2.6 QoS Classes and Parameters

In UNI 3.x the quality of service is indicated by a single parameter called "QoS Class," which is essentially an index to a network specific table of values for the actual QoS parameters. In TM/UNI 4.0 three QoS parameters may be individually signalled, and the signalled values override those implied by the QoS Class, which is still present. These parameters are the Cell Loss Ratio (CLR), Cell Transfer Delay (CTD), and Cell Delay Variation (CDV) [6].

A network provider may choose to associate other parameters, such as Severely Errored Cell Block Ratio, with a QoS Class definition, but these cannot be signalled individually. The ATM Forum UNI 3.0, 3.1 and TM 4.0 specs, following prior ITU specs, give vague qualitative definitions for QoS Classes 1 to 4. (QoS Class 0 is well-defined as "no QoS parameters defined".) Since our mapping is based on these specifications, we generally follow this guidance by setting the QoS Class value to 0 for UBR and ABR (as REQUIRED), 1 for CBR and rtVBR and 3 for nrtVBR. Note that the QoS Class follows the ATM service category, and not the IP service, to avoid combination that are unlikely to be supported. For example, if only nrtVBR is available for GS, then choosing QoS Class = 1 would probably result in connection failure. The QoS Class MUST NOT be interpreted as a way to add real-time behavior to an inherently non-real-time service.

The ITU has included a standard set of parameter values for a (small) number of QoS Classes in the latest version of Recommendation I.356 [21]. Network providers may choose to define further network-specific QoS Classes in addition to these. Note that the QoS class definitions in the new I.356 version might not align with the model we follow from the ATM Forum UNI specs. Apart from these definitions, there is no consistent agreement on QoS Class definitions among providers in practice.

The ATM QoS parameters have no explicitly signalled IP layer counterparts. The values that are signalled in the ATM network are determined by the IP service definitions and knowledge of the underlying ATM network characteristics, as explained below.

The ingress edge router SHOULD keep a table of QoS information for the set of egress routers that it may establish VCs with. This table may be simplified by using default values, but it will probably be good practice to maintain a table of current data for the most popular egress points. An edge device that initiates VC setup generally needs to have some way to propose initial value for CDV and CTD, even if they are changed by negotiation; so by positing such a table, we are not creating any new design burden. Cached information can be updated when VCs are successfully established, and to the extent that IP-layer reservations can wait for VCs to complete, the values can be refined through iterated negotiation.

Both GS and CLS REQUIRE that losses of packets due to congestion be minimized, so that the loss rate is approximately the same as for an unloaded network. The characteristic loss behavior of the physical medium not due to congestion (e.g., bit errors or fading on wireless channels) determines the order of the permitted packet loss rate. The ingress edge device MUST choose a value of CLR that provides the appropriate IP-level packet loss rate. The CLR value may be uniform

over all egress points in the ATM network, or may differ, e.g., when wireless or satellite ATM links are in some paths. The determination of CLR MUST account for the effects of packet size distribution and ATM Frame Discard mode (which can change the effective packet loss rate by orders of magnitude [22]).

The ingress router will also tabulate values for the Minimum Path Latency (MPL) and estimated queueing delays (D\_ATM) for each egress point. The latter will be used as part of the Adspec "D" parameter for GS, but its use here applies to CLS as well (when the VC setup includes delay parameters). MPL represents all constant (non-congestion related) delays, including propagation delay. D\_ATM accounts for the variable component of delays in the ATM network. (It may depend on non-signalled parameters such as CDVT.) Given these quantities, a new VC can be set up with delay-related QoS parameters given by

$$\begin{aligned}\text{CDV} &= \text{D\_ATM} \\ \text{CTD} &= \text{D\_ATM} + \text{MPL}.\end{aligned}$$

(CDV and CTD may be adjusted (increased) by the slack term in GS, see Section 3.3 below.)

It is interesting (and perhaps unfortunate) to note that in a typical GS/rtVBR service, the delay bound advertised can contain two components of  $b/R$  instead of one. Consider the simple case where  $\text{SCR} = R$  is the rate allocated to the flow in both IP routers and ATM switches along the path, and the buffer allocation is  $\text{MBS} = b$ . Parekh's theory [23], which is the basis of the GS delay formula [8] states that the  $b/R$  delay term occurs only once, because once a burst of size  $b$  has been served by a congested node at rate  $R$ , the packets will not arrive at a subsequent node as a single burst. However, we can't tell a priori if this bottleneck will occur in the ATM network or elsewhere in the IP network, so the declaration of CDV SHOULD account for it (i.e.,  $\text{CDV} \geq b/R$ ). Once CDV is set, the ATM network can impose this delay, whether or not the traffic arrives in a burst. Since the delay  $b/R$  can also occur elsewhere, it cannot be removed from the first term of the GS delay formula. The ATM  $b/R$  delay component appears in the third term of the GS delay formula,  $\text{D\_tot}$ . See Section 3.3 below for more on GS Adspec parameters. This effect may be mitigated when the ATM network employs more efficient statistical resource allocation schemes.

## 2.7 Additional Parameters -- Frame Discard Mode

TM/UNI 4.0 allows the user to choose a mode where the ATM network is aware, for the purpose of congestion management, of PDUs larger than an ATM cell (i.e., AAL PDUs that correspond in our context to IP packets). This facilitates implementation of algorithms such as partial packet discard, where a dropped cell causes subsequent cells in the same AAL-5 PDU to be dropped as well. Several other applicable buffer management schemes have been proposed [22, 24].

Frame discard can improve the efficiency and performance of end-to-end protocols such as TCP, since the remaining cells of a damaged PDU are generally useless to the receiver. For IP over ATM, Frame Discard MUST always be indicated, if available.

## 3.0 Additional IP-Integrated Services Protocol Features

### 3.1 Path Characterization Parameters for IP Integrated Services with ATM

This section discusses the setting of General Characterization Parameters (GCPs) at an ATM egress edge router. GCPs are signalled from IP source to IP destination, and modified by intermediate nodes using the Adspec portion of PATH messages in rsvp. The GS-specific Adspec parameters are discussed below in Section 3.3. These parameters are denoted as <x,y> where x is the service and y is the parameter number. Service number 1 indicates default or general parameter values. Please refer to [25] for definitions and details.

The IS break bit <1,2> MUST, of course, be left alone by implementations following these guidelines (as they are presumably IS-aware). Similarly, the router MUST always increment IS\_HOPS <1,4>. The GS and CLS service-specific break bits, <2,2> and <5,2> respectively, MUST be set if the support of the service is inadequate. In general GS is adequately supported by CBR (BCOB-A) and rtVBR service categories, and not adequately supported by UBR, ABR and nrtVBR because delays are not controlled. CLS may be adequately supported by all service categories except UBR (or Best Effort in UNI 3.x). See Sections 5, 6 for further discussion.

For GS, the ATM network MUST meet the delay performance advertised through the Adspec parameters, MPL, C, and D. If it cannot predictably meet these requirements, the GS break bit MUST be set. Similarly both break bits MUST be set if reservations are honored, but sufficient resources to avoid congestion loss are not allocated in practice. If the service break bits are not set, then the corresponding service hop counters, <2,4>, <5,4>, MUST be incremented.

The Available Path Bandwidth (APB) parameters <x,6> indicate the minimum physical bottleneck rate along the path. This may be discoverable in an ATM network as the negotiated PCR value for any UBR VC along the same path. This value MUST be corrected for AAL, ATM and physical-layer headers, as necessary, to reflect the effective IP datagram bandwidth. With ATM, it is possible that there is some policy limitation on the value of PCR, below the physical link bottleneck. In this case, the advertised value of APB (in general, and for each service if the values of APB signalled are service specific) MUST reflect this limit, since excess traffic beyond this rate will be dropped. (Note that there is no tagging of traffic in excess of PCR for TM/UNI 4.0.) These values SHOULD generally be cached by the ingress router for the set of egress routers with which it typically needs to establish VCs. The APB parameters are only adjusted down, to reflect the minimum as the composed value.

In the case of a multipoint VC, several parameters can be different for each egress point, e.g., because the characteristics of the physical links of the VC branches differ. When this occurs, the IWF at the egress routers MUST correct these values in PATH messages as they exit the ATM network. (We use the word "correct" because the ingress router SHOULD set the parameters to a value that is appropriate for the largest number of branches, or a value that would do the least harm if the egress routers failed to correct such parameters for each branch.) This is the only case where the egress router needs to operate on rsvp control messages. (A similar correction MUST be implemented for any non-rsvp set-up mechanism). The parameters for which such correction is REQUIRED are the Available Path Bandwidth (APB), the Minimum Path Latency (MPL), the Path MTU (although for ATM/AAL-5 this may typically be constant), and the ATM-specific components of the GS Adspec parameters C\_ATM and D\_ATM.

The ingress router table SHOULD store values for the ATM-network MPL <x,7> for the various egress points. The composed values <x,8> are formed by addition and forwarded along the path. In the cases where ATM routing chooses different paths, depending on the service category, for VCs to a given egress point, the table will generally reflect different values for each service. If the ATM network is very large and complex, it may become difficult to predict the routes that VCs will take once they are set up. This could be a significant source of misconfiguration, resulting in discrepancies between GS delay advertisements and actual results. The RSpec Slack term may be useful in mitigating this problem.

AAL-5 will support any message size up to 65,535 bytes, so setting the AAL SDU to the receiver TSpec M parameter value (plus 8 bytes for

the LLC/SNAP header) will generally not be an issue. In the PATH Adspec, however, the PATH\_MTU parameter <x,10> for each service SHOULD be set to 9188 bytes, which is the default MTU for AAL-5 [19].

### 3.2 Handling of Excess Traffic

For IP Integrated Services, network elements will transport traffic in excess of the TSpec parameters whenever physical resources (bandwidth, buffers and processing) are available. (In CLS a "network element MUST attempt to forward the excess traffic on a best-effort basis" under certain conditions; and in GS a traffic policers "SHOULD relegate non-conforming datagrams to best effort".) While excess traffic SHOULD be supported on a best effort basis, it MUST NOT interfere with the QoS (delay and loss) of conforming CLS and GS traffic, nor with normal service of non-reserved best effort traffic.

There are several solutions with ATM: the most attractive is to use a VBR service category (with an appropriate conformance definition) and tag excess traffic as low priority using the CLP bit. This avoids reordering of the flow, but necessitates careful design of the egress router scheduler. To avoid reordering, the excess traffic can be queued with conforming traffic. A threshold SHOULD be used to ensure that conforming traffic is not unnecessarily delayed by the excess. Also, for GS, the extra delay that would be incurred due to excess traffic in the queue ahead of conforming packets would have to be accurately reflected in the delay advertisement. Note that the ingress router SHOULD tag all cells of each non-conforming packet, rather than letting the ATM network apply tagging due to ATM-level non-conformance.

There is no requirement in ATM standards that tagged cells, marked either by the user or by policing, be transported if possible. Therefore, the operator of an edge router supporting IP-IS SHOULD ascertain the actual behavior of the ATM equipment in the path, which may span multiple administrative domains in the ATM network. If tagged cells are simply dropped at some point, regardless of load, then the operator may consider setting the break bit, at least for CLS service.

The other solutions generally involve a separate VC to carry the excess. A distinct VC can be set up for each VC supporting a GS or CLS flow, or, if many flows are aggregated into a single QoS VC, then another VC can handle the excess traffic for that set of flows. A VC can be set up to handle all excess traffic from the ingress router to the egress point. Since the QoS of the excess traffic is not particularly constrained, the design is quite flexible. However, using a separate VC may cause misordering of packets within a flow.



The service category for the excess-traffic VC may typically be UBR or ABR, although one could use CBR or nrtVBR if the excess traffic were predictable enough to know what rate to allocate. (This wouldn't normally be expected for excess traffic, though.)

Whether a separate VC is used may be influenced by the design of the router scheduler. The CLS spec suggests two possible implementations: one where excess traffic shares the Best Effort class scheduler allocation, but at lower priority than other best effort traffic. The other, where a separate allocation is made. The first would allow excess traffic to use the same VC as normal best effort traffic, and the second would suggest a separate VC.

TM/UNI 4.0. does not support tagging of traffic in excess of PCR. Although UNI 3.x does have a separate PCR parameter for CLP=0 cells only, we do not recommend using this feature for reasons of interoperability with TM/UNI 4.0 equipment. This restricts CBR VCs to use solutions other than tagging. The value of PCR can be set higher than necessary for conformant traffic, in an effort to support excess traffic on the same VC. In some cases this may be a viable solution, such as when there is little additional cost imposed for a high PCR. If PCR can be set as high as APB, then the excess traffic is fully accommodated.

### 3.3 Use of Guaranteed Service Adspec Parameters and Slack Term

The Adspec is used by the Guaranteed Service to allow a receiver to calculate the worst-case delay associated with a GS flow. Three quantities, C, D, and MPL, are accumulated (by simple addition of components corresponding to each network element) in the PATH message from source to receiver. The resulting delay values can be different for each unique receiver. The maximum delay is computed as

$$\text{delay} \leq b_r/R + C_{TOT}/R + D_{TOT} + MPL$$

The Minimum Path Latency (MPL) includes propagation delay, while  $b_r/R$  accounts for bursts due to the source and C and D include other queueing, scheduling and serialization delays. (We neglect the effect of maximum packet size and peak rate here; see the GS specification [8] for a more detailed equation.) The service rate requested by the receiver, R, can be greater than the TSpec rate,  $r_r$ , resulting in lower delay. The burst size,  $b_r$ , is the leaky bucket parameter from the receiver TSpec.

The values of C and D that a router advertises depend on both the router packet scheduler and the characteristics of the subnet attached to the router. Each router (or the source host) takes responsibility for its downstream subnet in its advertisement. For

example, if the subnet is a simple point-to-point link, the subnet-specific parts of C and D need to account for the link transmission rate and MTU. An ATM subnet is generally more complex.

For this discussion, we consider only the ATM subnet-specific components, denoted C\_ATM and D\_ATM. The ATM network can be represented as a "pure delay" element, where the variable queueing delay, given by CVD is captured in D\_ATM, and C\_ATM is set to zero. It is possible to use C\_ATM only when the ATM service rate equals R. This may be the case, for example with a CBR VC with PCR = R.

Usually it will be simpler to just advertise the total delay variation (CDV) in D\_ATM.

As discussed in Section 2.6, the edge router keeps a table with values of MPL and D\_ATM for each egress router it needs to share VCs with. The value of D\_ATM contributes to the D parameter advertised by the edge router, and SHOULD accurately reflect the CDV that the router will get in a VC when it is set up. Factors that affect CDV, such as statistical multiplexing in the ATM network, SHOULD be taken into account when compiling data for the router's table. In case of uncertainty, D\_ATM can be set to an upper bound. When an RESV message arrives, causing a VC to be set up, the requested values for CTD and CDV can be relaxed using the slack term in the receiver RSpec:

$$\begin{aligned} \text{CTD} &= \text{D\_ATM} + \text{MPL} + \text{S\_ATM} \\ \text{CDV} &= \text{D\_ATM} + \text{S\_ATM}. \end{aligned}$$

The term S\_ATM is the portion of the slack term applied to the ATM portion of the path. Recall that the slack term [8] is positive when the receiver can afford more delay than that computed from the Adspec. The ATM edge device may take part (or all) of the slack term, S. The distribution of delay slack among the nodes and subnets is network specific.

Note that with multipoint VCs the egress edge router may need to correct advertised values of C and D. See discussion in Section 3.1.

## 4.0 Miscellaneous Items

### 4.1 Units Conversion

All rates and token bucket depth parameters that are mapped from IP-level parameters to ATM parameters MUST be corrected for the effects of added headers and the segmentation of packets into cells. At the IP layer, token bucket depths and rates are measured in bytes and bytes/sec, respectively, whereas for ATM, they are measured in cells

and cells/sec.

Each IP Packet is wrapped into an AAL-5 PDU, having a number of additional header bytes (8 for LLC/SNAP and perhaps others, e.g. 12 for MPOA, etc.), and an 8-byte AAL-5 trailer. The AAL-5 PDU is then segmented into multiple ATM cells, each having a 5-byte cell header followed by a 48-byte cell payload. The number of cells used to carry an IP packet with

B = number of IP-packet Bytes,  
H = number of AAL-5 header bytes (LLC/SNAP etc.)  
C = number of cells,

is roughly

$C = B/48,$

and more precisely

$C = \text{floor}[(H + B + 8 + 47)/48]$

where floor[] is rounds down to the nearest integer. The '8' accounts for the AAL-5 trailer and the '47' accounts for the last cell which may be only partially filled.

## 5.0 Summary of ATM VC Setup Parameters for Guaranteed Service

This section describes how to create ATM VCs appropriately matched for Guaranteed Service. The key points are that real-time timing is REQUIRED, that the data flow may have a variable rate, and that demotion of non-conforming traffic to best effort is REQUIRED to be in agreement with the definition of GS. For this reason, we prefer an rtVBR service in which tagging is supported. Another good match is to use CBR with special handling of any non-conforming traffic, e.g., through another VC, since a CBR VC will not accommodate traffic in excess of PCR.

Note, these encodings assume point to multipoint connections, where the backward channel is not used. If the IP session is unicast only, then a point-to-point VC may be used and the IWF may make use of the backward channel, with QoS parameters set appropriately for the service provided.

We provide a mapping for all combinations of IP service and ATM service category, and comments indicating whether or not each combination meets the requirements of the IP-IS service.

## 5.1 Encoding GS Using Real-Time VBR (ATM Forum TM/UNI 4.0)

RtVBR with conformance definition VBR.3 [6] MEETS the requirements of GS.

## AAL

|                        |                                     |  |
|------------------------|-------------------------------------|--|
| Type                   | 5                                   |  |
| Forward CPCS-SDU Size  | parameter M of rcvr TSpec + 8 Bytes |  |
| Backward CPCS-SDU Size | parameter M of rcvr TSpec + 8 Bytes |  |
| SSCS Type              | 0 (Null SSCS)                       |  |

## Traffic Descriptor

|                            |                       |        |
|----------------------------|-----------------------|--------|
| Forward PCR CLP=0+1        |                       | Note 1 |
| Backward PCR CLP=0+1       | 0                     |        |
| Forward SCR CLP=0          |                       | Note 1 |
| Backward SCR CLP=0         | 0                     |        |
| Forward MBS (CLP=0)        |                       | Note 1 |
| Backward MBS (CLP=0)       | 0                     |        |
| BE indicator               | NOT included          |        |
| Forward Frame Discard bit  | 1                     |        |
| Backward Frame Discard bit | 1                     |        |
| Tagging Forward bit        | 1 (Tagging requested) |        |
| Tagging Backward bit       | 1 (Tagging requested) |        |

## Broadband Bearer Capability

|                          |                          |        |
|--------------------------|--------------------------|--------|
| Bearer Class             | 16 (BCOB-X)              | Note 2 |
| ATM Transfer Capability  | 9 (Real time VBR)        | Note 3 |
| Susceptible to Clipping  | 00 (Not Susceptible)     |        |
| User Plane Configuration | 01 (Point-to-Multipoint) |        |

## Broadband Low Layer Information

|                          |                      |        |
|--------------------------|----------------------|--------|
| User Information Layer 2 |                      |        |
| Protocol                 | 12 (ISO 8802/2)      |        |
| User Information Layer 3 |                      |        |
| Protocol                 | 11 (ISO/IEC TR 9577) | Note 4 |
| ISO/IEC TR 9577 IPI      | 204                  |        |

## QoS Class

|                    |   |        |
|--------------------|---|--------|
| QoS Class Forward  | 1 | Note 5 |
| QoS Class Backward | 1 | Note 5 |

## Extended QoS Parameters

|                        |  |        |
|------------------------|--|--------|
| Acceptable Forward CDV |  | Note 6 |
| Acceptable Forward CLR |  |        |
| Forward Max CTD        |  |        |

Note 1: See discussion in Section 2.5.1.

- Note 2: Value 3 (BCOB-C) can also be used.  
If Bearer Class C is chosen the ATC field MUST be absent.
- Note 3: The ATC value 19 is not used. The value 19 implies that the CLR objective applies to the aggregate CLP=0+1 stream and that does not give desirable treatment of excess traffic.
- Note 4: For QoS VCs supporting GS or CLS, the layer 3 protocol SHOULD be specified. For BE VCs, it can be left unspecified, allowing the VC to be shared by multiple protocols, following RFC 1755.
- Note 5: Cf ITU Rec. I.356 [21] for new QoS Class definitions.
- Note 6: See discussion in Section 2.6.

## 5.2 Encoding GS Using CBR (ATM Forum TM/UNI 4.0)

A CBR VC MEETS the requirements of GS. The main advantage of this is that CBR is widely supported; the disadvantage is that data flows might not fill the pipe (utilization loss) and there is no tagging option available. Excess traffic MUST be handled using a separate VC.

### AAL

|                        |                                     |  |
|------------------------|-------------------------------------|--|
| Type                   | 5                                   |  |
| Forward CPCS-SDU Size  | parameter M of rcvr TSpec + 8 Bytes |  |
| Backward CPCS-SDU Size | parameter M of rcvr TSpec + 8 Bytes |  |
| SSCS Type              | 0 (Null SSCS)                       |  |

### Traffic Descriptor

|                            |                           |        |
|----------------------------|---------------------------|--------|
| Forward PCR CLP=0+1        |                           | Note 1 |
| Backward PCR CLP=0+1       | 0                         |        |
| BE indicator               | NOT included              |        |
| Forward Frame Discard bit  | 1                         |        |
| Backward Frame Discard bit | 1                         |        |
| Tagging Forward bit        | 0 (Tagging not requested) |        |
| Tagging Backward bit       | 0 (Tagging not requested) |        |

### Broadband Bearer Capability

|                          |                          |        |
|--------------------------|--------------------------|--------|
| Bearer Class             | 16 (BCOB-X)              | Note 2 |
| ATM Transfer Capability  | 5 (CBR)                  | Note 3 |
| Susceptible to Clipping  | 00 (Not Susceptible)     |        |
| User Plane Configuration | 01 (Point-to-Multipoint) |        |

### Broadband Low Layer Information

|                          |                      |        |
|--------------------------|----------------------|--------|
| User Information Layer 2 |                      |        |
| Protocol                 | 12 (ISO 8802/2)      |        |
| User Information Layer 3 |                      |        |
| Protocol                 | 11 (ISO/IEC TR 9577) | Note 4 |
| ISO/IEC TR 9577 IPI      | 204                  |        |

## QoS Class

|                    |   |        |
|--------------------|---|--------|
| QoS Class Forward  | 1 | Note 5 |
| QoS Class Backward | 1 | Note 5 |

## Extended QoS Parameters

Note 6

Acceptable Forward CDV  
 Acceptable Forward CLR  
 Forward Max CTD

Note 1: See discussion in Section 2.5.1.

Note 2: Value 1 (BCOB-A) can also be used.

If Bearer Class A is chosen the ATC field MUST be absent.

Note 3: The ATC value 7 is not used. The value 7 implies CLR objective applies to the aggregate CLP=0+1 stream and that does not give desirable treatment of excess traffic.

Note 4: For QoS VCs supporting GS or CLS, the layer 3 protocol SHOULD be specified. For BE VCs, it can be left unspecified, allowing the VC to be shared by multiple protocols, following RFC 1755.

Note 5: Cf ITU Rec. I.356 [21] for new QoS Class definitions.

Note 6: See discussion in Section 2.6.

## 5.3 Encoding GS Using Non-Real-Time VBR (ATM Forum TM/UNI 4.0)

NrtVBR does not provide delay guarantees and is NOT RECOMMENDED for GS. If GS/nrtVBR is used and network utilization is low, the delay may be 'reasonable', but will not be controlled. The encoding of GS with nrtVBR is the same as that for CLS using nrtVBR. See Section 6.1 below.

## 5.4 Encoding GS Using ABR (ATM Forum TM/UNI 4.0)

GS using ABR is a very unlikely combination, and DOES NOT meet the service requirements of GS. The objective of the ABR service is to provide "low" loss rates. The delay objectives for ABR SHOULD be expected to be very loose. If ABR were used for GS, the VC parameters would follow as for CLS over ABR. See Section 6.2.

## 5.5 Encoding GS Using UBR (ATM Forum TM/UNI 4.0)

The UBR service is the lowest common denominator of the services. It cannot provide delay or loss guarantees, and therefore DOES NOT meet the requirements of GS. However if it is used for GS, it will be encoded in the same way as Best Effort over UBR, with the exception that the Forward PCR would be determined from the peak rate of the receiver TSpec. See Section 7.1.

## 5.6 Encoding GS Using ATM Forum UNI 3.0/3.1 Specifications

It is not recommended to support GS using UNI 3.x VBR mode because the BCOB-C Bearer Class does not represent real-time behavior. Also, Appendix F of the UNI 3.1 specification precludes the specification of traffic type "VBR" with the timing requirement "End to End timing Required" in conjunction with Bearer Class X.

A CBR VC MEETS the requirements of GS. The following table specifies the support of GS using CBR.

## AAL

|                        |                                     |        |
|------------------------|-------------------------------------|--------|
| Type                   | 5                                   |        |
| Forward CPCS-SDU Size  | parameter M of rcvr TSpec + 8 Bytes |        |
| Backward CPCS-SDU Size | parameter M of rcvr TSpec + 8 Bytes |        |
| Mode                   | 1 (Message mode)                    | Note 1 |
| SSCS Type              | 0 (Null SSCS)                       |        |

## Traffic Descriptor

|                      |                       |        |
|----------------------|-----------------------|--------|
| Forward PCR CLP=0    |                       | Note 2 |
| Backward PCR CLP=0   | 0                     |        |
| Forward PCR CLP=0+1  |                       | Note 2 |
| Backward PCR CLP=0+1 | 0                     |        |
| BE indicator         | NOT included          |        |
| Tagging Forward bit  | 1 (Tagging requested) |        |
| Tagging Backward bit | 1 (Tagging requested) |        |

## Broadband Bearer Capability

|                          |                          |        |
|--------------------------|--------------------------|--------|
| Bearer Class             | 16 (BCOB-X)              | Note 3 |
| Traffic Type             | 001 (Constant Bit Rate)  |        |
| Timing Requirements      | 01 (Timing Required)     |        |
| Susceptible to Clipping  | 00 (Not Susceptible)     |        |
| User Plane Configuration | 01 (Point-to-Multipoint) |        |

## Broadband Low Layer Information

|                          |                      |        |
|--------------------------|----------------------|--------|
| User Information Layer 2 |                      |        |
| Protocol                 | 12 (ISO 8802/2)      |        |
| User Information Layer 3 |                      |        |
| Protocol                 | 11 (ISO/IEC TR 9577) | Note 4 |
| ISO/IEC TR 9577 IPI      | 204                  |        |

## QoS Class

|                    |   |        |
|--------------------|---|--------|
| QoS Class Forward  | 1 | Note 5 |
| QoS Class Backward | 1 |        |

Note 1: Only included for UNI 3.0.

Note 2: See discussion in Section 2.5.1. PCR CLP=0 SHOULD be set

identical to PCR CLP=0+1. Although this could potentially allow a CBR VC to carry excess traffic as tagged cells, it is not recommended since it is not supported in UNI 4.0

Note 3: Value 1 (BCOB-A) can also be used. If BCOB-A is used Traffic Type and Timing Requirements fields are not included.

Note 4: For QoS VCs supporting GS or CLS, the layer 3 protocol SHOULD be specified. For BE VCs, it can be left unspecified, allowing the VC to be shared by multiple protocols, following RFC 1755.

Note 5: QoS Parameters are implied by the QoS Class.

## 6.0 Summary of ATM VC Setup Parameters for Controlled Load Service

This section describes how to create ATM VCs appropriately matched for Controlled Load Service. CLS traffic is partly delay tolerant and has variable rate. NrtVBR and ABR (TM/UNI 4.0 only) are the best choices for supporting CLS.

Note, these encodings assume point to multipoint connections where the backward channel is not used. If the IP session is unicast only, then a point-to-point VC may be used and the IWF may make use of the backward channel, with QoS parameters set appropriately for the service provided.

We provide a mapping for all combinations of IP service and ATM service category, and comments indicating whether or not each combination meets the requirements of the IP-IS service.

### 6.1 Encoding CLS Using Non-Real-Time VBR (ATM Forum TM/UNI 4.0)

NrtVBR MEETS the requirements for CLS.

#### AAL

|                        |                                     |
|------------------------|-------------------------------------|
| Type                   | 5                                   |
| Forward CPCS-SDU Size  | parameter M of rcvr TSpec + 8 Bytes |
| Backward CPCS-SDU Size | parameter M of rcvr TSpec + 8 Bytes |
| SSCS Type              | 0 (Null SSCS)                       |

#### Traffic Descriptor

|                            |              |        |
|----------------------------|--------------|--------|
| Forward PCR CLP=0+1        |              | Note 1 |
| Backward PCR CLP=0+1       | 0            |        |
| Forward SCR CLP=0          |              | Note 1 |
| Backward SCR CLP=0         | 0            |        |
| Forward MBS (CLP=0)        |              | Note 1 |
| Backward MBS (CLP=0)       | 0            |        |
| BE indicator               | NOT included |        |
| Forward Frame Discard bit  | 1            |        |
| Backward Frame Discard bit | 1            |        |



|                                 |                          |        |
|---------------------------------|--------------------------|--------|
| Tagging Forward bit             | 1 (Tagging requested)    |        |
| Tagging Backward bit            | 1 (Tagging requested)    |        |
| Broadband Bearer Capability     |                          |        |
| Bearer Class                    | 16 (BCOB-X)              | Note 2 |
| ATM Transfer Capability         | 10 (Non-real time VBR)   | Note 3 |
| Susceptible to Clipping         | 00 (Not Susceptible)     |        |
| User Plane Configuration        | 01 (Point-to-Multipoint) |        |
| Broadband Low Layer Information |                          |        |
| User Information Layer 2        |                          |        |
| Protocol                        | 12 (ISO 8802/2)          |        |
| User Information Layer 3        |                          |        |
| Protocol                        | 11 (ISO/IEC TR 9577)     | Note 4 |
| ISO/IEC TR 9577 IPI             | 204                      |        |
| QoS Class                       |                          |        |
| QoS Class Forward               | 3                        | Note 5 |
| QoS Class Backward              | 3                        | Note 5 |
| Extended QoS Parameters         |                          | Note 6 |
| Acceptable Forward CDV          |                          |        |
| Acceptable Forward CLR          |                          |        |
| Forward Max CTD                 |                          |        |

Note 1: See discussion in Section 2.5.2.

Note 2: Value 3 (BCOB-C) can also be used.

If Bearer Class C is used, the ATC field MUST be absent.

Note 3: The ATC value 11 is not used. The value 11 implies CLR objective applies to the aggregate CLP=0+1 stream and that does not give desirable treatment of excess traffic.

Note 4: For QoS VCs supporting GS or CLS, the layer 3 protocol SHOULD be specified. For BE VCs, it can be left unspecified, allowing the VC to be shared by multiple protocols, following RFC 1755.

Note 5: Cf ITU Rec. I.356 [21] for new QoS Class definitions.

Note 6: See discussion in Section 2.6.

## 6.2 Encoding CLS Using ABR (ATM Forum TM/UNI 4.0)

ABR MEETS the requirements for CLS when MCR is set to the CLS TSpec rate.

### AAL

|                        |                                     |
|------------------------|-------------------------------------|
| Type                   | 5                                   |
| Forward CPCS-SDU Size  | parameter M of rcvr TSpec + 8 Bytes |
| Backward CPCS-SDU Size | parameter M of rcvr TSpec + 8 Bytes |

|                                 |                           |        |
|---------------------------------|---------------------------|--------|
| SSCS Type                       | 0 (Null SSCS)             |        |
| Traffic Descriptor              |                           |        |
| Forward PCR CLP=0+1             |                           | Note 1 |
| Backward PCR CLP=0+1            | 0                         |        |
| Forward MCR CLP=0+1             |                           | Note 1 |
| Backward MCR CLP=0+1            | 0                         |        |
| BE indicator                    | NOT included              |        |
| Forward Frame Discard bit       | 1                         |        |
| Backward Frame Discard bit      | 1                         |        |
| Tagging Forward bit             | 0 (Tagging not requested) |        |
| Tagging Backward bit            | 0 (Tagging not requested) |        |
| Broadband Bearer Capability     |                           |        |
| Bearer Class                    | 16 (BCOB-X)               | Note 2 |
| ATM Transfer Capability         | 12 (ABR)                  |        |
| Susceptible to Clipping         | 00 (Not Susceptible)      |        |
| User Plane Configuration        | 00 (Point-to-Point)       |        |
| Broadband Low Layer Information |                           |        |
| User Information Layer 2        |                           |        |
| Protocol                        | 12 (ISO 8802/2)           |        |
| User Information Layer 3        |                           |        |
| Protocol                        | 11 (ISO/IEC TR 9577)      | Note 3 |
| ISO/IEC TR 9577 IPI             | 204                       |        |
| QoS Class                       |                           |        |
| QoS Class Forward               | 0                         | Note 4 |
| QoS Class Backward              | 0                         | Note 4 |
| Extended QoS Parameters         |                           | Note 5 |
| Acceptable Forward CDV          |                           |        |
| Acceptable Forward CLR          |                           |        |
| Forward Max CTD                 |                           |        |
| ABR Setup Parameters            |                           | Note 6 |
| ABR Additional Parameters       |                           | Note 6 |

Note 1: See discussion in Section 2.5.2.

Note 2: Value 3 (BCOB-C) can also be used.

If Bearer Class C is chosen the ATC field MUST be absent.

Note 3: For QoS VCs supporting GS or CLS, the layer 3 protocol SHOULD be specified. For BE VCs, it can be left unspecified, allowing the VC to be shared by multiple protocols, following RFC 1755.

Note 4: Cf ITU Rec. I.356 [21] for new QoS Class definitions.

Note 5: See discussion in Section 2.6.

Note 6: The ABR-specific parameters are beyond the scope of this document. These generally depend on local implementation and not on values mapped from IP level service parameters (except for MCR). See [6, 11] for further information.

### 6.3 Encoding CLS Using CBR (ATM Forum TM/UNI 4.0)

Although CBR does not explicitly take into account the variable rate of source data, it may be convenient to use ATM connectivity between edge routers to provide a simple "pipe" service, as a leased line replacement. Since no tagging option is available with CBR, excess traffic **MUST** be handled using a separate VC. Under this condition, CBR **MEETS** the requirements of CLS.

To use CBR for CLS, the same encoding for GS over CBR (Section 5.2) would be used. See discussion in Section 2.5.2.

### 6.4 Encoding CLS Using Real-Time VBR (ATM Forum TM/UNI 4.0)

The encoding of CLS using rtVBR implies a hard limit on the end-to-end delay in the ATM network. This creates more complexity in the VC setup than the CLS service requires, and is therefore not a preferred combination, although it **DOES MEET** the requirements of CLS.

If rtVBR is used to encode CLS, then the encoding is essentially the same as that for GS. See discussions in Section 5.1 and Section 2.5.2.

### 6.5 Encoding CLS Using UBR (ATM Forum TM/UNI 4.0)

This encoding gives no QoS guarantees and **DOES NOT MEET** the requirements of CLS. If used, it is coded in the same way as for BE over UBR (Section 7.1), except that the PCR would be determined from the peak rate of the receiver TSpec.

### 6.6 Encoding CLS Using ATM Forum UNI 3.0/3.1 Specifications

This encoding is equivalent to the nrtVBR service category. It **MEETS** the requirements of CLS.

#### AAL

|                        |  |
|------------------------|--|
| Type                   | 5  |
| Forward CPCS-SDU Size  | parameter M of rcvr TSpec + 8 Bytes          |
| Backward CPCS-SDU Size | parameter M of rcvr TSpec + 8 Bytes          |
| Mode                   | 1 (Message mode)                      Note 1 |
| SSCS Type              | 0 (Null SSCS)                                |

## Traffic Descriptor

|                      |                       |        |
|----------------------|-----------------------|--------|
| Forward PCR CLP=0+1  |                       | Note 2 |
| Backward PCR CLP=0+1 | 0                     |        |
| Forward SCR CLP=0    |                       | Note 2 |
| Backward SCR CLP=0   | 0                     |        |
| Forward MBS (CLP=0)  |                       | Note 2 |
| Backward MBS (CLP=0) | 0                     |        |
| BE indicator         | NOT included          |        |
| Tagging Forward bit  | 1 (Tagging requested) |        |
| Tagging Backward bit | 1 (Tagging requested) |        |

## Broadband Bearer Capability

|                          |                          |        |
|--------------------------|--------------------------|--------|
| Bearer Class             | 16 (BCOB-X)              | Note 3 |
| Traffic Type             | 010 (Variable Bit Rate)  |        |
| Timing Requirements      | 00 (No Indication)       |        |
| Susceptible to Clipping  | 00 (Not Susceptible)     |        |
| User Plane Configuration | 01 (Point-to-Multipoint) |        |

## Broadband Low Layer Information

|                          |                      |        |
|--------------------------|----------------------|--------|
| User Information Layer 2 |                      |        |
| Protocol                 | 12 (ISO 8802/2)      |        |
| User Information Layer 3 |                      |        |
| Protocol                 | 11 (ISO/IEC TR 9577) | Note 4 |
| ISO/IEC TR 9577 IPI      | 204                  |        |

## QoS Class

|                    |   |        |
|--------------------|---|--------|
| QoS Class Forward  | 3 | Note 5 |
| QoS Class Backward | 3 | Note 5 |

Note 1: Only included for UNI 3.0.

Note 2: See discussion in Section 2.5.2.

Note 3: Value 3 (BCOB-C) can also be used. If BCOB-C is used Traffic Type and Timing Requirements fields are not included.

Note 4: For QoS VCs supporting GS or CLS, the layer 3 protocol SHOULD be specified. For BE VCs, it can be left unspecified, allowing the VC to be shared by multiple protocols, following RFC 1755.

Note 5: Cf ITU Rec. I.356 [21] for new QoS Class definitions. QoS Parameters are implied by the QoS Class.

## 7.0 Summary of ATM VC Setup Parameters for Best Effort Service

This section is provided for completeness only. The IETF ION working group documents on ATM signalling support for IP over ATM [10, 11] provide definitive specifications for Best Effort IP service over ATM.

The best-matched ATM service category to IP Best Effort is UBR. We provide the setup details for this case below. The BE service does not involve reservation of resources. ABR and nrtVBR are also well suited to BE service. See discussion in Section 2.1.3.

Note, VCs supporting best effort service are usually point to point, rather than point to multipoint, and the backward channels of VCs are used. In cases where VCs are set up to support best effort multicast sessions, multipoint VCs can be used and the backward channels would be not have resources reserved. Related situations include transport of excess traffic on IP-multicast QoS sessions, or to support the subset of multicast end systems that have not made rsvp reservations. See the discussion on VC management in [12].

## 7.1 Encoding Best Effort Service Using UBR (ATM Forum TM/UNI 4.0)

### AAL

|                        |                                    |  |
|------------------------|------------------------------------|--|
| Type                   | 5                                  |  |
| Forward CPCS-SDU Size  | 9188 Bytes (default MTU for AAL-5) |  |
| Backward CPCS-SDU Size | 9188 Bytes (default MTU for AAL-5) |  |
| SSCS Type              | 0 (Null SSCS)                      |  |

### Traffic Descriptor

|                            |                       |        |
|----------------------------|-----------------------|--------|
| Forward PCR CLP=0+1        |                       | Note 1 |
| Backward PCR CLP=0+1       | 0                     |        |
| BE indicator               | included              |        |
| Forward Frame Discard bit  | 1                     |        |
| Backward Frame Discard bit | 1                     |        |
| Tagging Forward bit        | 1 (Tagging requested) |        |
| Tagging Backward bit       | 1 (Tagging requested) |        |

### Broadband Bearer Capability

|                          |                          |        |
|--------------------------|--------------------------|--------|
| Bearer Class             | 16 (BCOB-X)              | Note 2 |
| ATM Transfer Capability  | 10 (Non-real time VBR)   |        |
| Susceptible to Clipping  | 00 (Not Susceptible)     |        |
| User Plane Configuration | 01 (Point-to-Multipoint) |        |

### Broadband Low Layer Information

|                          |                 |        |
|--------------------------|-----------------|--------|
| User Information Layer 2 |                 |        |
| Protocol                 | 12 (ISO 8802/2) | Note 3 |

### QoS Class

|                    |   |
|--------------------|---|
| QoS Class Forward  | 0 |
| QoS Class Backward | 0 |

Note 1: See discussion in Section 2.5.3.

Note 2: Value 3 (BCOB-C) can also be used.

If Bearer Class C is used, the ATC field MUST be absent

Note 3: For QoS VCs supporting GS or CLS, the layer 3 protocol SHOULD be specified. For BE VCs, it can be left unspecified, allowing the VC to be shared by multiple protocols, following RFC 1755.

## 8.0 Security Considerations

IP Integrated Services (including rsvp) and ATM are both complex resource reservation protocols, and SHOULD be expected to have complex feature interactions.

Differences in IP and ATM billing styles could cause unforeseen problems since RESV messages can set up VCs. For example, an end-user paying a flat rate for (non-rsvp aware) internet service may send an rsvp RESV message that encounters a (perhaps distant) ATM network with a usage-sensitive billing model. Insufficient authentication could result in services being accidentally billed to an innocent third party, intentional theft of service, or malicious denial of service attacks where high volumes of reservations consume transport or processing resources at the edge devices.

The difference in styles of handling excess traffic could result in denial of service attacks where the ATM network uses transport resources (bandwidth, buffers) or connection processing resources (switch processor cycles) in an attempt to accommodate excess traffic that was admitted by the internet service.

Problems associated with translation of resource reservations at edge devices are probably more complex and susceptible to abuse when the IP-ATM edge is also an administrative boundary between service providers. Note also that administrative boundaries can exist within the ATM cloud, i.e., the ingress and egress edge devices are operated by different service providers.

Note, the ATM Forum Security Working Group is currently defining ATM-level security features such as data encryption and signalling authentication. See also the security issues raised in the rsvp specification [3].

## 9.0 Acknowledgements

The authors received much useful input from the members of the ISSLL working group. In particular, thanks to Drew Perkins and Jon Bennett of Fore Systems, Roch Guerin of IBM, Susan Thomson and Sudha Ramesh of Bellcore.

## Appendix 1 Abbreviations

|              |   |
|--------------|---|
| AAL          | ATM Adaptation Layer                            |
| ABR          | Available Bit Rate                              |
| APB          | Available Path Bandwidth (int-serv GCP)         |
| ATC          | ATM Transfer Capability                         |
| ATM          | Asynchronous Transfer Mode                      |
| B-LLI        | Broadband Low Layer Information                 |
| BCOB         | Broadband Connection-Oriented Bearer Capability |
| BCOB-{A,C,X} | Bearer Class A, C, or X                         |
| BE           | Best Effort                                     |
| BT           | Burst Tolerance                                 |
| CBR          | Constant Bit Rate                               |
| CDV          | Cell Delay Variation                            |
| CDVT         | Cell Delay Variation Tolerance                  |
| CLP          | Cell Loss Priority (bit)                        |
| CLR          | Cell Loss Ratio                                 |
| CLS          | Controlled Load Service                         |
| CPCS         | Common Part Convergence Sublayer                |
| CTD          | Cell Transfer Delay                             |
| EOM          | End of Message                                  |
| GCP          | General Characterization Parameter              |
| GCRA         | Generic Cell Rate Algorithm                     |
| GS           | Guaranteed Service                              |
| IE           | Information Element                             |
| IETF         | Internet Engineering Task Force                 |
| ION          | IP Over Non-broadcast multiple access networks  |
| IP           | Internet Protocol                               |
| IPI          | Initial Protocol Identifier                     |
| IS           | Integrated Services                             |
| ISSLL        | Integrated Services over Specific Link Layers   |
| ITU          | International Telecommunication Union           |
| IWF          | Interworking Function                           |
| LIJ          | Leaf Initiated Join                             |
| LLC          | Logical Link Control                            |
| MBS          | Maximum Burst Size                              |
| MCR          | Minimum Cell Rate                               |
| MPL          | Minimum Path Latency                            |
| MTU          | Maximum Transfer Unit                           |
| nrtVBR       | Non-real-time VBR                               |
| PCR          | Peak Cell Rate                                  |
| PDU          | Protocol Data Unit                              |
| PVC          | Permanent Virtual Connection                    |
| QoS          | Quality of Service                              |
| RESV         | Reservation Message (of rsvp protocol)          |
| RFC          | Request for Comments                            |
| RSVP         | Resource Reservation Protocol                   |
| RSpec        | Reservation Specification                       |

|       |   |
|-------|---|
| rtVBR | Real-time VBR   |
| SCR   | Sustainable Cell Rate                                   |
| SDU   | Service Data Unit                                       |
| SNAP  | Subnetwork Attachment Point                             |
| SSCS  | Service-Specific Convergence Sub-layer                  |
| SVC   | Switched Virtual Connection                             |
| TCP   | Transport Control Protocol                              |
| TM    | Traffic Management                                      |
| TSpec | Traffic Specification                                   |
| UBR   | Unspecified Bit Rate                                    |
| UNI   | User-Network Interface                                  |
| UPC   | Usage Parameter Control (ATM traffic policing function) |
| VBR   | Variable Bit Rate                                       |
| VC    | (ATM) Virtual Connection                                |

## REFERENCES

- [1] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [2] Braden, R., Clark, D., and S. Shenker, "Integrated Services in the Internet Architecture: an Overview", RFC 1633, June 1994.
- [3] Braden, R., Zhang, L., Berson, S., Herzog, S., and S. Jamin, "Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification", RFC 2205, September 1997.
- [4] The ATM Forum, "ATM User-Network Interface Specification, Version 3.0", Prentice Hall, Englewood Cliffs NJ, 1993.
- [5] The ATM Forum, "ATM User-Network Interface Specification, Version 3.1", Prentice Hall, Upper Saddle River NJ, 1995.
- [6] The ATM Forum, "ATM User-Network Interface (UNI) Signalling Specification, Version 4.0", July 1996. Available at <ftp://ftp.atmforum.com/pub/approved-specs/af-sig-0061.000.ps>.
- [7] The ATM Forum, "ATM Traffic Management Specification, Version 4.0", April 1996. Available at <ftp://ftp.atmforum.com/pub/approved-specs/af-tm-0056.000.ps>.
- [8] M. W. Garrett, "A Service Architecture for ATM: From Applications to Scheduling", IEEE Network Mag., Vol. 10, No. 3, pp. 6-14, May 1996.
- [9] Shenker, S., Partridge, C., and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, September 1997.



- [10] Wroclawski, J., "Specification of the Controlled-Load Network Element Service", RFC 2211, September 1997.
- [11] Perez, M., Liaw, F., Mankin, A., Hoffman, E., Grossman, D., and A. Malis, "ATM Signaling Support for IP over ATM", RFC 1755, February 1995.
- [12] Maher, M., "ATM Signalling Support for IP over ATM - UNI Signalling 4.0 Update", RFC 2331, April 1998.
- [13] Crawley, E., Berger, L., Berson, S., Baker, F., Borden, M., and J. Krawczyk, "A Framework for Integrated Services and RSVP over ATM", RFC 2382, August 1998.
- [14] Berger, L., "RSVP over ATM Implementation Requirements", RFC 2380, August 1998.
- [15] Berger, L., "RSVP over ATM Implementation Guidelines", BCP 24, RFC 2379, August 1998.
- [16] Shenker, S., and J. Wroclawski, "Network Element Service Specification Template", RFC 2216, September 1997.
- [17] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", RFC 2210, September 1997.
- [18] Borden, M., Crawley, E., Davie, B., and S. Batsell, "Integration of Real-time Services in an IP-ATM Network Architecture", RFC 1821, August 1995.
- [19] Heinanen, J., "Multiprotocol Encapsulation over ATM Adaptation Layer 5", RFC 1483, July 1993.
- [20] Laubach, M., "Classical IP and ARP over ATM", RFC 1577, January 1994.
- [21] ITU Recommendation I.356, "B-ISDN ATM layer cell transfer performance", International Telecommunication Union, Geneva, October 1996.
- [22] A. Romanow, S. Floyd, "Dynamics of TCP Traffic over ATM Networks", IEEE J. Sel. Areas in Commun., Vol. 13, No. 4, pp. 633-41, May 1995.

- [23] A. K. Parekh, R. G. Gallager, "A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Multiple Node Case", IEEE/ACM Trans. Networking, Vol. 2, No. 2, pp. 137-150, April 1994.
- [24] S. Floyd, V. Jacobson, "Link-sharing and Resource Management Models for Packet Networks", IEEE/ACM Trans. Networking, Vol. 3, No. 4, August 1995.
- [25] S. Shenker and J. Wroclawski, "General Characterization Parameters for Integrated Service Network Elements", RFC 2215, September 1997.

#### Authors' Addresses

Mark W. Garrett  
Bellcore  
445 South Street  
Morristown, NJ 07960  
USA

Phone: +1 201 829-4439  
EMail: mwg@bellcore.com

Marty Borden  
Bay Networks  
42 Nagog Park  
Acton MA, 01720  
USA

Phone: +1 508 266-1011  
EMail: mborden@baynetworks.com

## Full Copyright Statement

Copyright (C) The Internet Society (1998). All Rights Reserved.

This document and translations of it may be copied and furnished to others, and derivative works that comment on or otherwise explain it or assist in its implementation may be prepared, copied, published and distributed, in whole or in part, without restriction of any kind, provided that the above copyright notice and this paragraph are included on all such copies and derivative works. However, this document itself may not be modified in any way, such as by removing the copyright notice or references to the Internet Society or other Internet organizations, except as needed for the purpose of developing Internet standards in which case the procedures for copyrights defined in the Internet Standards process must be followed, or as required to translate it into languages other than English.

The limited permissions granted above are perpetual and will not be revoked by the Internet Society or its successors or assigns.

This document and the information contained herein is provided on an "AS IS" basis and THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIMS ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

