Compressed Video Over Networks

Chapter 4 ATM Networks

Arthur W. Berger
Lucent Technologies
1600 Massachusetts Ave. #707
Cambridge MA 02138-2923, U.S.A.
phone & fax: +1 617 547 9352
e-mail: awberger@lucent.com

4.1 Introduction

Asynchronous Transfer Mode (ATM) is a technology and suite of standards for high-speed packet networks that uses small, fixed-size packets called cells and provides connection-oriented transfer and a variety of service classes. ATM networks can provide high-speed transmission with low cell loss and low delay variation and thus are well suited for supporting the transport of video. This chapter provides an overview of ATM. Basic concepts are covered, with greater detail given to aspects of most relevance for the transport of video.

As shown in Figure 1 the ATM cell is 53 bytes in length and consists of 5 bytes of header followed by 48 bytes of payload into which user-information can be placed. ATM is connection oriented in the sense that a connection is first established between a source and destination, and then ATM cells are relayed along the connection, arriving in order at the destination.

Figure 1. ATM Cell

As in any connection-oriented packet network, the ATM connection is "virtual" in the sense that a particular portion of a transport facility, such as a time-slot, is not dedicated to the given connection, which, in contrast, would have been the case in a "circuit-switched" connection such as in traditional telephony networks. Various types of connections can be established, providing a variety of service classes, some more suitable for the transport of video than others, as will be discussed in Section 4.7. The virtual connections can be established by signaling protocols for on-demand transport, or, on a slower time scale, can be established by management plane procedures, [1 - 3]. The establishment by signaling protocols is analogous to the dialing of a telephone call, while establishment via management plane procedures is analogous to private-line service.[*]

The ATM standards, such as those for signaling, are in recommendations issued by the International Telecommunications Union (ITU) and specifications issued by the ATM Forum. The recommendations of the ITU are approved after consensus is reached by the delegates from participating countries. This process can result in recommendations that lack the specificity that would be desirable for simple, interoperable implementations. Beginning in 1991, the ATM Forum, an industry forum of member companies, began preparing specifications that were based on existing and draft ITU recommendations and that were focused on unambiguous implementations. The ATM Forum progressed its work more quickly than the ITU and began to

---

[*] From the viewpoint of supporting the transport of video, the mechanism by which an ATM connection is established is not a major concern, and thus due to limited space is not covered in this overview. The reader interested in signaling and associated topics of addressing and routing can consult the references discussed in Section 4.9, as well as [1 –3].

lead on some topics. To the extent that the ATM Forum and the ITU worked on common topics, individual participants endeavored obtain consistency between the two bodies, and were largely, but not completely, successful in doing so. Except where noted below, the ATM standards reviewed in this chapter are consistent between these two bodies. As a practical matter, one can consider that both the ITU and the ATM Forum specify the standards for ATM, and the distinction between a recommendation versus a specification is of little significance.

The ATM standards partition functions of an ATM network into three layers: the physical layer, (and with some redundancy of terminology) the ATM layer, and the ATM Adaptation layer. At end-points of the ATM connection all three layers are present, while at intermediate nodes, there are just the first two layers, as illustrated in Figure 2. The end-point of the ATM connection can be the actual end-point of the user information transfer, such as a personal computer, or a video terminal. Alternatively, the ATM connection may only be a portion of the end-to-end transfer of the user information. As an example, the ATM connection may span only the backbone of a given service provider, in which case the end-point of the ATM connection would be at a node (a gateway, router, or switch) on the path of the information transfer. In either case, at the end-point of the ATM connection, the user information is in some non-ATM format, such as a file stored in memory, or an Internet Protocol (IP) packet. The ATM Adaptation layer packages, adapts, the higher-layer information into 48 byte chunks. This packaged user information is placed into an ATM cell at the ATM layer, which also associates the cell with a particular ATM connection. The ATM cell is then processed by the physical layer and transmitted onto the physical transmission path. At intermediate nodes along the ATM connection, the ATM cell header needs to be examined to determine the appropriate egress port from the node (the appropriate next transmission path). At the receiving end-point of the ATM connection, the user information is then extracted from the ATM cell and returned to the higher-layer format.

Figure 2. Protocol stack in ATM networks

ATM connections, particularly those that pass through a public network, are associated with a traffic contract. A traffic contract is part of the service agreement with the public network provider whereby the provider commits to meeting specified quality of service objectives, given that the ATM cell flow entering the provider's network is in conformance with the given type of ATM connection. The present chapter, in the context of an overall introduction to ATM, gives relatively greater weight to describing these quality of service aspects and the types of ATM connections and discusses their applicability for the transport of video.

4.1.1 Outline of Remainder of Chapter.

The Physical, ATM, and Adaptation layers are briefly described in Sections 4.2, 4.3, and 4.4, respectively, with emphasis on the ATM layer. Section 4.5 then reviews the quality of service that can be provided by ATM networks. Section 4.6 presents the different types of ATM connections, known as "service categories" or "ATM transfer capabilities." Section 4.7 discusses the strengths and weakness of the different ATM connections from the viewpoint of transporting encoded video, and Section 4.8 briefly discusses the deployment of ATM in core and access networks. Section 4.9 points the reader to further readings, and Section 4.10 lists the acronyms used in the chapter.

4.2 ATM Physical Layer

The ATM Physical Layer provides the needed functionality to insert and extract ATM cells onto and from physical transmission media, such as fiber, coaxial cable, and twisted copper wires. The ATM Physical Layer (abbreviated as PHY) has been defined for various interfaces, and Table 1 provides a sampling.

Table 1. Example physical interfaces for which an ATM Physical Layer is defined.

The ATM Physical Layer is divided into two sublayers: the Transmission Convergence Sublayer, and the Physical Media Dependent Sublayer, where the latter is of less interest for present purposes and, as the name suggests, is concerned with functionality that is dependent on the particular media. ITU Recommendation I.432, [4], contains details on the functions of the physical layer, and Onvural, [5], provides a more readable presentation, as well as more details than given here.

4.2.1 Transmission Convergence Sublayer of the ATM Physical Layer

The Transmission Convergence Sublayer accepts cells from the ATM Layer and forms a bit stream for the Physical Media Dependent Sublayer, and likewise, accepts a bit stream from the Physical Media Dependent Sublayer and passes cells to the ATM Layer. Some of the functions of the Transmission Convergence Sublayer are: Header Error Control, framing, cell delineation, cell-rate decoupling, and cell scrambling.

Header Error Control (HEC) provides protection from bit errors in the ATM cell header that can occur, primarily during the propagation of the cell between ATM nodes. The HEC can detect and correct one-bit errors and can detect some multiple bit errors. At the transmitting end, the ATM Layer passes to the PHY layer the first four bytes of the cell header, and the 48-byte payload. The Transmission Convergence Sublayer then computes the HEC value, which is then placed in the fifth byte of the cell header. At the receiving end, the Transmission Convergence Sublayer checks the integrity of the cell header, and if no errors are detected, then passes the ATM cell to the ATM layer.

Framing is used by some physical layers, such as Synchronous Optical Network (SONET) and Synchronous Digital Hierarchy (SDH). A frame consists of overhead bytes and a payload, which of present interest would support the transfer of ATM cells. In the case of SONET and SDH a frame is generated synchronously, one frame each 125 microseconds, and the overhead bytes provide a rich functionality for operations, including restoration from network failures, administration, and maintenance. The Transmission Convergence Sublayer is responsible for the generating the appropriate framing structure and for the placement of cells within it.

Cell delineation determines the location of the boundary of an ATM cell, from amongst the stream of bits from the physical medium. The receiving end of the ATM connection is the "hunt state" when it is searching for the cell boundary, which occurs at connection startup, or when the cell boundary is lost, for example, due to a burst error somewhere along the transmission path. A basic search mechanism is to continuously compute the HEC value on 32 contiguous bits and see if it matches with the next 8 bits, and if a match occurs, then a candidate cell header, and hence cell boundary, is found, and the receiver moves to the "pre-synch state." If the subsequent *n* presumed cell headers also check out, then the receiver presumes that the cell boundary is indeed detected, and declares itself to be in the "synch state." Depending on the physical medium, the byte boundaries may already be known. Also, some framing structures contain a pointer that identifies the location of the cell boundary.

Cell rate decoupling provides a continuous stream of cells to the physical interface, regardless of whether the ATM layer has anything to send. Cell rate decoupling is analogous to the continuous transmission of bits, regardless of whether higher layer information is being conveyed. Most of the physical interfaces in Table 1 require that a continuous stream of cells be sent. When no cells are being handed down from the ATM layer, the Transmission Convergence Sublayer itself generates cells, called idle cells. Idle cells have a particular cell header, not used by any other cell, and their the payload contains a particular byte repeated 48 times.

Scrambling ensures that the bit stream contains sufficiently frequent changes in the bit value, so that timing and synchronization can be maintained by the Physical Medium-Dependent Sublayer. Scrambling avoids the situation of the transmission of many contiguous bits with the same value. Another technique is block coding at the Physical Medium-Dependent Sublayer, whereby a group of 4 bits, say, is encoded as a 5-bit block.

4.2.2 Physical Medium-Dependent Sublayer of the ATM Physical Layer

Some of the functions of the Physical Medium-Dependent Sublayer are: (1) encoding of bits for transmission, (2) timing and synchronization, and (3) the actual insertion and extraction of the electrical or optical signal to and from the physical medium.

4.3 ATM Layer

4.3.1 ATM cell header

The ATM cell header is 5 bytes in length, and has two, slightly different, standardized formats depending on the interface. One interface is called the user-network interface (UNI) and is the interface between a public network and an ATM end-system, or a public network and a private (enterprise or corporate) network. The other interface is called the network-network interface (NNI) (a.k.a. as the network-node interface within the ITU) and is the interface between two public networks or between two private networks, or between two switches within a public or private network. At the user-network interface, the ATM cell header consists of the following fields:

1. Generic Flow Control (GFC), 4 bits in length,
2. Virtual Path Identifier (VPI), 8 bits,
3. Virtual Channel Identifier (VCI), 12 bits
4. Payload Type (PT), 3 bits
5. Cell Loss Priority (CLP), 1 bit,
6. Header Error Control (HEC), 8 bits.

The position of the fields in the cell header is shown in Figure 3.

Figure 3. ATM cell header at the User-Network Interface (UNI)[*]

The ATM cell header at the Network-Node Interface (NNI) is the same as at the UNI, except the GFC field is not present, and instead the VPI field is extended from 8 bits to 12 bits.

The **Generic Flow Control** (GFC) field is generally not used and the field is set equal to zero. The intent of GFC was for media access control for ATM cells on bus, ring, or star topology customer-premise networks, which are on the "user" side of the user-network interface, and where the network element on "network" side of the UNI had primary control. The GFC protocol procedures are defined in ITU Recommendation I.361, [6].

The **Virtual Path Identifier** (VPI) and the **Virtual Channel Identifier** (VCI) associate the ATM cell with a given connection, and enable the multiplexing of many connections on a single transmission path. The VPI and VCI fields are discussed in Section 4.3.2 below. The lowest values of the VPI and VCI field are reserved for special functions, such as signalling, and operations, administration and maintenance, see ITU Recommendation I.361 for details, [6].

As the name suggests, the **Payload Type** (PT) field identifies the type of payload in the ATM cell, as well as some additional indications. The 3 bits of the payload type field allow for 8 code points, whose assigned meaning is summarized in Table 2.

---

[*] Material from ITU recommendations is reproduced with prior authorization of the ITU as copyright holder. Complete recommendations can be obtained from: ITU, Place des Nations, CH-1211 Geneva 20, Switzerland. http://www.itu.int/publications.

Table 2. Standardized code points in the Payload Type field.[*]

As shown in Table 2, the first four code points are for user-data cells, the next two code points identify cells for Operations, Administration and Maintenance (OAM), see Recommendation I.610, [7], the seventh code point identifies the resource management cell, which is used for feedback control schemes to mediate contention for resources, and the last code point is reserved. The first four code points distinguish user-data cells with respect to two independent attributes. The first attribute is the binary "ATM-user-to-ATM-user" indication, where "ATM-user" refers to the protocol layer that is above ATM layer, namely the ATM adaptation layer, (AAL). In particular, this indication is used by AAL type 5 to denote the last cell of a given AAL5 protocol data unit, see Section 4.4.3. The second attribute indicates whether the given cell passed through a network element that is or may soon be in congestion. User-data cells are originally transmitted with the PT code point equal to (in binary) 000, or 001. An intermediate network element that is in congestion may change the second bit of the PT field from 0 to 1. This function is called Explicit Forward Congestion Indication. See ITU Recommendation I.371 for details, [8].

The **Cell Loss Priority** (CLP) bit can be used to distinguish between cells of a given ATM connection with regard to loss priority. Depending on network conditions, cells with the CLP bit set equal to 1 (CLP=1) are subject to discard prior to cells with the CLP bit set equal to 0 (CLP=0), [8]. The CLP bit is used, particularly, in the ATM service category called "variable bit rate," as described in Section 4.6.2 below.

---

[*] Source: Ref. [6]

4.3.2 ATM Connections: VPCs, VCCs, Label Swapping

The term "ATM connection" applies to both Virtual Channel Connections (VCCs) and Virtual Path Connections (VPCs). In informal usage, "ATM connection" is used for either the uni-directional connection (which supports cells from a given source to a given destination), or the bi-directional connection, which consists of the pair of uni-directional connections, one for each direction of communication. Formally, the term ATM connection pertains to the uni-directional connection. However, the uni-directional connections are always established in pairs, even if the user-information is passing only in one direction, as OAM support for the ATM connection requires two-way communication.

ATM has a two-level hierarchy of ATM connections, whereby a group of VCCs is supported by a VPC. Figure 4 illustrates this hierarchy. Within a transmission path between two ATM nodes, there can be multiple VPCs, and within each VPC there can be multiple VCCs. (This hierarchy does not have to be used, and a VPC can contain a single VCC, in which case the end-points of the VPC would coincide with the end-points of the VCC.)

Figure 4. Transmission Path supports multiple VPCs, which in turn support multiple VCCs.

This hierarchy has been found to be very useful. (As discussed in Chapter 3 of the present volume, MultiProtocol Label Switching for IP networks extends the idea of a hierarchy to an arbitrary number of levels). A network service provider may wish to establish VPCs that extend across its network and which aggregate VCCs that pass through that provider's network. The network service provider could aggregate based on the service category of the VCC, see Section

4.6. Also, the service provider may be renting capacity to other carriers, in which case the VPC's could be dedicated to the support of a given carrier's VCCs. Likewise, a corporation might establish VPCs through a network service provider to interconnect various building locations of the corporation. The presence of VPCs allows for quicker establishment of signalled (switched) VCCs. The VPC's also enable faster restoration. Although a VPC is often chosen to be a constant rate "pipe", a VPC is not restricted to be so, and could be any of the service categories discussed in Section 4.6. For more details on VPCs see ITU Recommendation E.735, [9].

The Virtual Channel Identifier (VCI) and the Virtual Path Identifier (VPI) in the ATM cell header provide the information needed by ATM nodes to identify the VCC or VPC to which the given cell belongs. For a given ATM connection, the value of the VPI/VCI does not remain constant on all of the links the connection passes through. The VPI/VCI is not a destination address. Rather, for a given ATM connection, the value of the VPI/VCI pertains for a given link between two ATM nodes. An intermediate ATM node upon receiving a cell can determine from the VPI/VCI the appropriate egress port for this cell and can determine the appropriate value for the VPI/VCI that identifies the given ATM connection on the determined egress link. The ATM node maintains a table that maps the VPI/VCI of an incoming cell to the VPI/VCI to be inserted in the header of the cell upon egress. (With the change in VPI/VCI value, the header error control is recomputed, Section 4.2.1.) Thus, the value of the VPI/VCI has local significance, local to a given transmission path between two ATM nodes. The operation of switching the value of the VPI/VCI at intermediate nodes is called "label switching," or "label swapping," and is also being used in MultiProtocol Label Switching for IP networks.

Figure 5 illustrates label swapping in the context of ATM. Figure 5 shows a VCC originating at end-system A and terminating at end-system B, and passing through three intermediate network nodes. Between network nodes 1 and 3, the VCC is carried inside of a VPC. The particular values for the VPIs and VCIs in the figure were chosen arbitrarily, and are of interest only to the extent that they are different on different portions of the connections.

Figure 5. Illustration of label swapping of VPI/VCI values for a given Virtual Channel Connection.

Between end-system A and network node 1, cells belonging to the given VCC have a VPI=71 and a VCI=63. For each cell that network node 1 receives on a given input port, it examines the value of the VPI/VCI to determine the egress port. In the present example, a cell with VPI/VCI equal to 71/63 is to be forwarded to a network VPC that runs between network nodes 1 and 3. At the egress port of network node 1, this VPC is identified by VPI=37, and the cells of the given connection are identified by VCI=56. Thus, network node 1 replaces, the VPI/VCI value of 71/63 with 37/56. Network node 1 could be transmitting cells of many VCCs on this given VPC. As the length of the VCI field is 16 bits, in principle the number of VCCs that this VPC could support is 2^{16} (minus 32, as VCI values 0 through 31 are pre-assigned or reserved for future functions, [6]).

When network node 2 receives a cell on the given input port with VPI=37, network node 2 knows that this VPI is associated with a VPC that continues on to network node 3. Thus, network node 2 does not need to look at or modify the VCI's of cells of this VPC, as network node 2 does not need to be concerned with the individual VCC's within this VPC. Network node 2 replaces the VPI value of 37 with the value 96 when it transmits cells of this VPC to network node 3. As network node 2 does not change the VCI of any of the VCCs within this VPC, the VCI of the given connection remains at 56.

Network node 3 knows that arriving cells with VPI=96 on the given input port are associated with a VPC that ends at network node 3. Thus, network node 3 also examines the VCI, and from its forwarding table, knows that a cell with VPI/VCI = 96/56 should egress on a particular port and have the VPI/VCI values replaced with 82/43. End-system B on receipt of cells with VPI/VCI = 82/43 knows that these are cells from End-system A.

Quite likely end-system B is transmitting cells to end-system A on the companion, paired, VCC that runs from B to A. Even if the user data is going in just one direction, control traffic will be present on the reverse connection, such as OAM cells, or cells transporting Transmission Control Protocol (TCP) acknowledgements. In order that each network node can easily associate a VCC with its companion paired VCC, the same values for the VPI and VCI are used, at each segment along the path, for the cells travelling in each direction on this pair of VCCs.

4.3.3 ATM Connections: Point-to-Multipoint

The ATM connections discussed so far have been point-to-point, i.e. they have a single source and a single destination. The ATM layer also supports point-to-multipoint connections, which have a single source and multiple destinations. The source end emits one copy of each cell, and each cell is then replicated at nodes within the network. A natural application for point-to-multipoint connections is for video distribution to multiple customers. Point-to-multipoint connections are also used in the emulation of Local Area Networks (LANs). The ATM Forum has specified a LAN emulation protocol, [10], which includes a multicast server, called a Broadcast and Unknown Server (BUS). The BUS is connected to all of the LAN emulation clients via a point-to-multipoint VCC, and each LAN emulation client is connected to the BUS via a point-to-point VCC. If a LAN emulation client wishes to send a higher layer packet, such as an IP packet, to all of the clients, it sends the packet to the BUS, which in turns forwards it to all of the clients via the point-to-multipoint VCC. (The BUS needs to transmit all of cells associated with a given packet before sending any cells of another packet.) There are many more aspects to LAN Emulation, see [10] for details. The ATM Forum has also specified a generalization of LAN emulation that operates across the wide area, across multiple logical IP subnets. This extension is call MultiProtocol Over ATM (MPOA), [11].

4.4 ATM Adaptation Layer (AAL)

As illustrated in Figure 2, ATM Adaptation Layer (AAL) is above the ATM layer and below any higher layers. At the transmitting end of an ATM connection, the AAL takes higher layer Protocol Data Units (PDUs), such as IP packets, and segments them into ATM cells, and at the receiving end, does the reverse function of reassembly. In addition, the AAL provides functionality that enhances the service provided by the ATM layer. This additional functionality is tailored to support various classes of service or types of traffic. Thus, multiple AALs have been defined, each tailored to support particular services.

The following sections provide an overview of the different types AALs. For all of the AALs, the functions are divided into sublayers. Standardization has focused on the sublayers designated the Segmentation and Reassembly (SAR) sublayer, and the Common Part of the Convergence Sublayer (CPCS).

4.4.1 AAL Type 1

AAL Type 1 (AAL1) supports circuit emulation. As specified in ITU Recommendation I.363.1, [12], the functionality provided by AAL 1 includes:

- segmentation and reassembly of higher-layer information;
- handling of cell delay variation;
- handling of cell payload assembly delay;
- handling of lost or misinserted cells;
- source clock frequency recovery at the receiver;
- recovery of the source data structure at the receiver.

To provide these functions, AAL1 uses one byte of the 48 byte payload. Thus, 47 bytes per ATM cell are available for the transport of higher-layer information. The AAL1 convergence sublayer provides two methods for the support of asynchronous constant-bit-rate services where the clocks are not locked to a network clock. The two methods are "Adaptive Clock" and "Synchronous Residual Time Stamp," [12, Sections 2.5.2.2 and 2.5.1.2].

4.4.2 AAL Type 2

AAL Type 2 (AAL2) is a recently standardized AAL designed for low-rate, delay sensitive applications that generate short, variable-length packets. A motivating application for AAL2 was low-bit-rate voice, where the delay to fill the payload of an ATM cell with the encoded speech from a single voice source would have degraded performance. Thus, a key attribute of AAL2 is the ability to multiplex higher-layer streams onto a single ATM virtual channel connection (VCC). Up to 255 higher-layer streams can be multiplexed on a single VCC, which is possible since AAL2 does not require that each encapsulated packet fit within a single ATM payload, but rather can span across payloads. Another key attribute is the ability to handle higher-layer packets of variable-length, which occur from various encoding schemes. AAL2 is specified in ITU Recommendation I.363.2 [13]. A more readable description is provided by Baldwin et al. in [14]; and Sriram and Wang analyze the performance of AAL2 in the context of supporting voice with bit dropping, [15].

Although the specification of AAL Type 2 was motivated by voice, the specification does not restrict AAL2 to that application. AAL2 is also suitable to encapsulate low-bit-rate (less than 64Kbps), video streams, such as H.263 or MPEG-4 video encodings which can have rates in the range of 20kbps, or video encodings originally design for transport over the current public switched telephone network.

4.4.3 AAL Type 5

AAL Type 5 (AAL5) is designed for variable rate applications that do not require a timing or synchronization between the source and destination, or if the application does need such synchronization, then the application can obtain it by means other than the AAL.

AAL5 supports the non-assured transfer of higher-layer PDUs, where the length of the PDU can be from 1 to 65,535 bytes. AAL5 maintains the sequence of the user data, and can detect transmission errors, [16]. As discussed in Section 4.3.1, the "ATM user to ATM user indication" in the ATM cell header distinguishes the last ATM cell that segments a given higher-layer PDU.

The overhead for AAL5 (other than the "ATM user to ATM user indication") is placed in the rear of the last ATM cell that segments a given higher-layer PDU. The overhead consists of 8 bytes plus any padding that may be needed to fill out the payload of the cell. Thus, for example, a 1,024 byte higher-layer PDU would be segmented into 22 cells, where the full payload of the first 21 cells is used for the higher-layer PDU, and the 22nd cell contains the remaining 16 bytes of the PDU, followed by 24 bytes of padding, and finally 8 bytes of AAL5 protocol overhead, called the "Common Part Convergence Sublayer (CPCS) – PDU trailer." This trailer includes a 32 bit cyclic redundancy check and a 16 bit length field, which is used to determine the end of the user-data PDU (and the start of any padding) and also to detect any lost or misinserted cells.

4.4.4 AAL Type 3/4

AAL Type 3/4 (AAL3/4) was originally defined as two separate AALs: AAL3 was designed for connection-oriented variable-rate services, while AAL4 was designed for connectionless (i.e. datagram) variable-rate services. It was found that these two AALs could be combined into a single, harmonized AAL, called AAL3/4. AAL3/4 provides significantly more functionality than AAL5; in particular AAL3/4 can multiplex higher-layer flows onto a single ATM connection, but at the expense of greater complexity and of a greater overheard – four bytes out of each 48 byte payload. Shortly before the completion of AAL3/4, AAL5 was proposed and has subsequently gained wide acceptance as providing a better trade-off of simplicity versus features.

4.5 Quality of Service

An important feature of ATM networks is the performance commitments that can be provided to connections. The performance commitments can be at the connection level, such as the delay in establishing the connection or the probability a request for connection establishment is denied. ITU recommendation I.358 [17] covers performance at the connection level. In addition, and the focus of greater interest, are performance commitments at the cell level, such as cell loss and delay. The following sub-section summarizes cell-level performance parameters that have been defined. And then Section 4.5.2 reviews the associated performance objectives.

4.5.1 Quality of Service and Network Performance Parameters

ITU Recommendation I.356, [18], defines the following cell-level network performance parameters:

1. Cell Transfer Delay (CTD),
2. Cell Delay Variation (CDV),
3. Cell Loss Ratio (CLR),
4. Cell Error Ratio,
5. Cell Misinsertion Rate,
6. Severely Errored Cell Block Ratio.

The ATM Forum adopts the same parameters and uses the term "Quality of Service" (QoS) parameters, [19]. The following are informal definitions of the above parameters; precise definitions are given in [18] and [19].

Cell transfer delay (CTD) is the delay of a cell between to two reference locations, such as the end-points of an ATM connection, or the ingress and egress of a service provider's network. ITU defines the mean transfer delay as the arithmetic average of a specified number of cell transfer delays. The performance objective is in terms of an upper bound on the mean CTD. The ATM Forum defines the maximum cell transfer delay, denoted maxCTD, as a high quantile on the histogram of transfer delays; that is, the probability that the cell transfer delay is greater than this quantile is small.

Heuristically, cell delay variation is the variation in the delay. ITU defines two types of cell delay variation – a "one-point" and "two-point." The one-point cell delay variation is defined at a single reference location (and thus is relatively easy to measure) and quantifies the variability in the cell arrival times at the reference location with respect to equally-spaced, nominal reference times. This notion of delay variation is useful for measuring deviation from the peak cell rate in constant-bit-rate connections, see Section 4.6.1 below. The two-point cell delay variation is defined with respect to two reference locations and is the difference between the realized CTD and a reference cell transfer delay. The ATM Forum defines a third notion, called the peak-to-peak cell delay variation, which equals the maxCTD minus the fixed portion of the delay, and thus is approximately the difference between the shortest and longest cell transfer delays. ATM Forum's peak-to-peak CDV is essentially the range of the distribution of ITU's two-point CDV, if the reference delay of the latter is taken to be the fixed delay.

The Cell Loss Ratio (CLR) is the ratio of the total number of lost cells to the total number of transmitted cells for a population of interest, such as all of the cells transmitted on a connection, or on multiple instances of connections with a given end-point. The cell loss ratio excludes cells in "severely errored cell blocks," see below, though this may not be followed in some service offerings.

An errored cell is a cell that, although received at a reference location, has its binary content in error. The cell error ratio is the ratio of errored cells to received cells.

A misinserted cell is a cell that is received at a downstream reference location on a given connection but was never present on the connection at an upstream reference location. A misinserted cell can occur if bit errors corrupt the VPI/VCI of a cell of a given connection, and the corrupted VPI/VCI happens to be a valid VPI/VCI of another ATM connection sharing the transmission path. In such a case, the first connection experiences a cell loss, and the second connection experiences a misinserted cell. The cell misinsertation rate is the number of minsinserted cells observed over a specified time interval.

A cell block is "N" consecutively transmitted cells of a given connection. A severely errored cell block is a cell block where more than "M" errored, lost, or misinserted cells have occurred. The severely errored cell block ratio is the ratio of severely errored cell blocks to the total cell blocks for a population of interest.

4.5.2 Quality of Service Objectives

ITU Recommendation I.356, [18], includes provisional QoS classes and associated performance objectives. There are currently four QoS classes, labeled: stringent class, tolerant class, bi-level class, and "U" class, where the "U" stands for unbounded or unspecified. The QoS class can be selected on a connection by connection basis.

The associated performance objectives are stated for a very stressful scenario of a 27,500 kilometer reference connection with 25 nodes and one geo-stationary satellite hop. Thus, for "typical" connections, the performance objectives may appear rather weak. In practice, a service provider offering ATM service would state their own performance objectives.

Three of the QoS parameters, the cell error ratio, the cell misinsertion rate, and the severely errored cell block ratio, are primarily caused by impairments in the physical transmission media and are thus not readily controlled, or selectable, on a connection by connection basis (modulo selecting or avoiding routes with particular transmission media). Thus, these parameters are given default values that apply across all classes (except the "U" class for which no performance objectives are stated). The default values are as follows. The upper bound on the cell error ratio is $4*10^{-6}$. The upper bound on the mean cell misinsertion ratio is 1/day. The upper bound on the severely errored cell block ratio is 10^{-4}. For each of these objectives, as well as those in the sequel, explanatory notes are provided in Recommendation I.356, [18].

The remaining three QoS parameters, cell transfer delay, cell delay variation, and cell loss ratio, are influenced by the choice of route and by buffering and scheduling capabilities and policies in network nodes, as well as influenced by physical transmission media. Thus, it is reasonable for a user to be able to select on a per-connection basis objectives for these QoS parameters, from amongst the objectives supported by the network service provider. In particular, the user can select on a per-connection basis the QoS class, where each class is associated with a unique set of objective values. Table 3 summarizes the provisional performance objectives in ITU Recommendation I.356 associated with each QoS class (and where the entry "U" stands for unspecified, or unbounded). Again, see Recommendation I.356 for further detail and explanation of these objectives.

Table 3. Provisional Performance Objectives in ITU Recommendation I.356.

4.5.3 Associating QoS objectives to a given ATM connection

Applications that are using an ATM network, such as for the transport of video, can require certain quality of service from the network. In such cases, when the ATM connection is established, the user application would wish to have certain quality of service objectives pertain for that connection. As discussed in Section 4.5.2, QoS parameters that are primarily determined by physical layer impairments, such as cell error ratio, have default values that can not be selected on a per-connection basis. However, the QoS parameters of cell transfer delay, cell delay variation, and cell loss ratio are affected by the ATM layer, and it is reasonable to select objectives for these QoS parameters on a per-connection basis.

One way for the user application to select the quality of service objectives is to choose one of the QoS classes offered by the service provider. This method is supported by the signaling procedures defined by the ITU and the ATM Forum, [1, 2, 3].

A second way for the user application to select the quality of service objectives is to signal an acceptable value for the cell transfer delay, the delay variation, and/or the cell loss ratio. The user application could do this without needing to know the particulars of the QoS classes offered by the service provider. If the network, or networks, through which the connection would pass can meet the objectives signalled by the user, then the connection would be established, and otherwise it would be denied, in which case the application could possibly try another carrier, or could modify the requested performance. The ATM Forum UNI and Private Network-Network Interface specifications [2, 3] support this procedure for the above three QoS parameters, and the ITU Recommendation Q.2931, [1], supports this procedure for cell transfer delay. For cell transfer delay and cell delay variation, the signalling message also includes a field for the cumulative impairment. The value in this field is increased as the setup message passes from network to network, or node to node, including the portion within the end-user's network. If the cumulative value becomes greater than the acceptable value, then the connection is denied. The user application can also indicate that any value is acceptable (i.e. and thus not a priori impose a limit), and then can learn from the cumulative value what will be the likely performance.

Note that the QoS objectives are probabilistic, and need not hold for each realization of a connection. For example, if the cell loss ratio objective is 10^{-6}, then heuristically at least 10^{8}

cells should be transported to obtain a reasonable measurement, but this is more than a Gigabyte of user data and thus could be more than the content transmitted. Service providers state the performance objectives in the context of averaging over multiple connections or over a period of time.

Requesting particular QoS for a connection makes sense for certain types of ATM connections. In particular setting delay and loss objectives is most pertinent for ATM connections that are "constant bit rate" or "real-time variable bit rate." Both of these types of ATM connections are tailored to support the transport of video. These and other types of ATM connections are the subject of Section 4.6.

4.5.4 Quality of Service and Video

The above material is from the perspective of QoS at the ATM layer. For the transport of video, the QoS that the user cares about is the perceived quality of the decoded and displayed video. Unfortunately the relationship between the QoS at the ATM layer and the user's perceived quality of the decoded video is not well understood. However, one can make some fairly obvious, qualitative statements.

The end-to-end cell transfer delay objective is important for interactive video applications, where the delay experienced by the user also includes the additional delays in the video end-systems. Subjective tests on interactive voice communication have shown that if the delay experienced by the user is within 100ms, then the interaction feels normal, i.e. the delay is not noticed as an impairment. If the delay reaches 150ms, then some impairment is noticed, at least for a proportion of test subjects. When the delay is as big as 250ms, which occurs with a geo-stationary satellite, then the interaction can feel quite awkward, particularly for people who tend to start talking before the current talker completes the sentence. In contrast, for non-interactive video applications, such as playback of stored video, or live broadcast of a sports event or the nightly news, then the end-to-end delay objective is not important.

Cell losses cause the loss of a higher layer protocol data unit, such as a transport packet. Thus, cell losses are bad, they will occur, and a natural question is: at what level of cell loss can the decoder still obtain video of good quality? Unfortunately, the question does not have a simple answer, as many factors are pertinent. A cyclic dependency occurs where networking people want to be told by video people "the cell loss objective that the network needs to meet," so as to design and engineer the network appropriately. And the video people want to be told by the networking people "the cell loss objective that the network does meet," so as to design video encoding and decoding algorithms that work well given this level of cell loss. One reference point is that the best, i.e. smallest, cell loss ratio that the ATM network can meet is around 10^{-9} to 10^{-10}. This can occur for constant-bit-rate connections (see Section 4.6) where cell loss only occurs due to physical layer impairments. Although constant-bit-rate connections are appropriate in some cases, they are not the complete solution since they are relatively expensive (they inefficiently use network bandwidth) for video traffic that is actually variable rate, where the constant rate of the connection is set equal to peak rate of the video traffic. Thus, additional solutions are sought that make better use of the statistical multiplexing potential of ATM networks and of the inherent variable rate of encoded video, which brings us back to the above cycle. There seems to be some convergence, where cell loss rates in range of 10^{-4} to 10^{-6} are often mentioned. One potentially complicating factor is that cell losses tend to occur in bunches. Thus for a given objective of say 10^{-5}, cell losses tend not to be isolated at roughly one out of every 100,000 cells. Rather, ten consecutive cells might be lost, with longer intervals of no loss.

This bunching of cell losses occurs for an aggregate cell stream, and an individual VCC may indeed experience only isolated cell losses, depending on circumstances and implementation in the network node. Another factor is that the impact of cell loss on the displayed video depends on attributes of the encoding. If there is no inter-frame dependency, then the impact is confined to a single frame, while with inter-frame dependency (which has the advantage of enabling significant compression), the impact of cell losses can be quite noticeable.

To deal with cell losses, a possible variation is to make use of the cell-loss-priority (CLP) bit in the ATM cell header. Here the less significant, lower order, encoded bits would be transmitted in the lower priority CLP=1 cells, while the more significant, higher order encoded bits would be transmitted in the higher priority CLP=0 cells. If the network needs to discard a cell due to congestion at a buffer in a network node, then the lower priority CLP=1 cells would be discarded first. Although this provides more flexibility, cell losses of higher priority CLP=0 cells can still occur, and one would still need to establish a cell loss objective for the CLP=0 cells.

Providing objectives on cell delay variation within the ATM network is useful for sizing the build-out delay buffers at the receiving end. The cells that arrive late are equivalent to cells that are lost. Thus, it is sensible to have the performance objective on cell delay variation be as strict as the objective on cell loss (and in the case of layered coding, as strict as the loss objective for the high priority, CLP=0, cells).

4.6 ATM Service Categories (a.k.a. ATM Transfer Capabilities)

The ITU and the ATM Forum have defined roughly the same service categories, though with different terminology, which is a source of confusion. As both terminologies are used, it is worthwhile to be aware of them. Thus, in this paragraph we mention both terminologies, and thereafter use the vocabulary from the ATM Forum, as it is somewhat more popular. The term "service category" is used by the ATM Forum, while the ITU uses the term "ATM transfer capability." ITU Recommendation I.371, [8], defines four ATM transfer capabilities: (1) Deterministic Bit Rate, (2) Statistical Bit Rate, (3) Available Bit Rate, and (4) ATM Block Transfer. The ATM Forum in the Traffic Management Specification [19] also defines four service categories. The first three are the same as in Recommendation I.371, except the ATM Forum uses the term “Constant Bit Rate” instead of “Deterministic Bit Rate,” and the term “Variable Bit Rate” instead of “Statistical Bit Rate.” The fourth service category defined by the ATM Forum is called Unspecified Bit Rate. These terms are summarized in Table 4.

Table 4. Comparison of vocabulary in the ITU and ATM Forum.

Although there is some overlap in the features of the service categories and more than one may be suitable for a given application, the set of all service categories is intended to be able to support effectively a wide range of applications. In addition, both the ITU and the ATM Forum are studying additional service categories to supplement the ones already defined. For example, the ATM Forum and the ITU are expected to complete the definition of the Guaranteed Frame Rate service category, which is an enhancement to Unspecified Bit Rate, and provides a minimum cell rate commitment to conforming blocks of cells, called frames, [20-21].

The following sub-sections provide an overview of the above five service categories. Then Section 4.7 discusses the suitability of the various the service categories for transporting video.

4.6.1 Constant Bit Rate

The Constant Bit Rate (CBR) service category is used by connections that request a static volume of bandwidth, which is then available at any time during the lifetime of the connection. The volume of bandwidth is specified by the Peak Cell Rate (PCR).

In practice, the specification of a rate at a location in the network, such as at an interface, needs a second parameter in order to be well defined. Without this second parameter, it is unclear how strict or loose is the intended meaning. For example, a very strict interpretation would apply to every inter-cell time, and thus not allow for any delay variation that might occur in upstream nodes, or (for an absurd example) would prevent the specification any peak rate between 150Mbps and 75Mbps on an transmission path of capacity 150Mbps, since to attain such rates a portion of the cells have to be placed in adjacent cell-slot times, i.e. back-to-back and thus at a rate of 150Mbps. At the other extreme, a very loose interpretation would allow a source to request a peak cell rate of 10Mbps, and then if the source is idle for a minute, it could then transmit at 20Mbps for the next minute. Thus, a second parameter is needed to provide a tolerance, or an averaging, over some time scale. The PCR is specified with a tolerance, known as the Cell Delay Variation (CDV) tolerance. The specification is made precise via an algorithm called the Generic Cell Rate Algorithm (GCRA), where the PCR and the CDV tolerance are parameters in the algorithm.

The GCRA is shown in Figure 6 and is specified by both the ITU and ATM Forum, respectively in [8] and [19]. The GCRA has two parameters, the increment, denoted "I" and the limit, denoted "L." When the GCRA is used to define the notion of peak cell rate, then the increment I is set equal to the reciprocal of the peak cell rate, and the limit L is set equal to the CDV tolerance. The GCRA is another name for the leaky bucket algorithm and for the virtual scheduling algorithm. The later two, although different in appearance, are isomorphic with regard to the key attribute that for any arriving stream of cells, both algorithms will detect the same cells to be conforming, and consequently the same cells to be non-conforming.

Figure 6. Generic Cell Rate Algorithm (From Ref. [8].)

Often in the literature the "token bank" algorithm is used synonymously with the leaky-bucket algorithm. However, these two algorithms are not isomorphic in sense above, of detecting the same set of cells to be conforming. Though, for most practical purposes they are close enough that the differences are negligible, see Berger and Whitt [22, Section 7]. Moreover, a fluid version of the token-bank algorithm is isomorphic to the leaky bucket, where conceptually instead of discrete tokens arriving to the bank, a continuous flow of credit arrives. The GCRA has advantages over another popular algorithm known as the sliding window, [23]; and for a broader perspective on desirable properties for algorithms for traffic descriptors, see Berger [24], and Berger and Eckberg [25].

In the CBR service category the source can emit cells at the PCR, or any lower rate, at any time and for any duration, and the QoS commitments still pertain. A natural QoS to select for a CBR connection is the "stringent" QoS class specified in ITU Recommendation I.356 and summarized in Section 4.5.2 above. A natural service agreement by a network operator would be that if the submitted traffic is conforming to the specified PCR, then the network operator commits to a QoS that includes a specified cell-loss-ratio objective and an end-to-end cell-delay-variation objective sufficient to support an application relying on circuit emulation.

An obvious application for the CBR service category is to support circuit emulation at a higher layer. However, this is not the only application. With circuit emulation, from the viewpoint of the ATM layer, there is an ongoing stream of cells, nominally spaced at the reciprocal of the PCR. In contrast, another type of ATM connection that the CBR service category supports is one where the source may emit cells at any rate less than or equal to the PCR, and may emit no cells for periods. An important example is semi-permanent connections (leased lines) that may be established for periods of months but whose usage, at the ATM layer, varies during the course of the day. An important special case is a semi-permanent user-to-user Virtual Path Connection that is established between two locations of a corporation.

The traffic controls used in the CBR service category include the negotiation of the PCR and the subsequent policing of the submitted cell flow for conformance to the PCR. To obtain the QoS commitment of a tight end-to-end cell delay variation, the network operator needs to isolate the CBR connections from other types of connections. One way network nodes attain this isolation is to serve the CBR connections with higher delay priority than other types of connections.

4.6.2 Variable Bit Rate

In the Variable Bit Rate (VBR) service category, the end-system uses standardized traffic parameters to specify, in greater detail than just the Peak Cell Rate, the cell flow that will be emitted on the connection. The standardized traffic parameters are: the Peak Cell Rate (PCR), the Sustainable Cell Rate (SCR), and the Maximum Burst Size (MBS). (In addition, associated tolerances are used at public interfaces.) The source may also describe the traffic via the declaration of the “Service Type” that pertains for the connection [8]. Informally, the SCR is thought of as the average rate of a connection. More precisely, the SCR is an upper bound on the average cell rate of the connection, where average cell rate is the total number of cells transmitted divided by the duration of the connection. In the VBR service category, the SCR is always strictly less than the PCR, and often is between a third and a tenth of the PCR. The Maximum Burst Size (MBS) is the maximum number of cells that may be sent at the Peak Cell Rate and still be conforming to the traffic descriptor.

Conformance to the pair of parameters SCR and MBS is made precise via the Generic Cell Rate Algorithm (GCRA), see Figure 6. The increment I in the GCRA is set equal to the reciprocal of the SCR. The limit L is determined from the three parameters PCR, SCR, and MBS and is set equal to L = [MBS – 1][(1/SCR) – (1/PCR)], see [19, section C.4] for more details. In the context of SCR, the limit L is called the Burst Tolerance, a.k.a. the Intrinsic Burst Tolerance. The MBS, in a somewhat convoluted way, is the second parameter that provides the needed tolerance for the SCR. The traffic parameters are typically static for the duration of the connection, though their values can be re-negotiated via signalling or management procedures.

The VBR service category makes use of the Cell Loss Priority (CLP) bit in the ATM cell header. The CLP bit distinguishes between high priority (CLP=0) cells and low priority (CLP=1) cells of a connection. The CLP bit can be set by the source, or, at the request of the user, the network can use the “tagging” feature, whereby for cells submitted with CLP=0 that are not conforming to the SCR/MBS traffic parameters for CLP=0 cells, the network changes the CLP bit to 1. In the VBR service category, three configurations of the traffic parameters along with the CLP bit and the tagging option are specified:

1. The PCR and the SCR/MBS traffic parameters apply to the aggregate of all cells, i.e. regardless of the value of the CLP.
2. The PCR applies to the aggregate of all cells, and the SCR/MBS traffic parameters apply only to CLP=0 cells.

3. The same as (2) above except with the tagging option invoked.

When the connection has both CLP=0 and CLP=1 cells, the cells entering the network with CLP=0 can be viewed as the committed portion of the traffic (the QoS commitment includes a specified cell loss ratio for the CLP=0 cells) and the cells entering the network with CLP=1 (or cells tagged by the network as CLP=1) can be viewed as the "at risk portion" (the cell loss ratio is Unspecified for the CLP=1 cells).

To precisely define the above three configurations, again the GCRA was used – twice, once for the PCR and once for the SCR/MBS. There are multiple ways that the two GCRAs can put together, all a priori reasonable, but unfortunately inconsistent. This caused some confusion. One way was chosen: cells are first tested with respect to the PCR, and then tested with respected to the SCR/MBS, where the specifics of the second test depend on which of the three configurations pertains. Figure 7 presents the algorithm for the second configuration, i.e. the SCR/MBS only applies to the CLP=0 cells and tagging is not used. The algorithm for the first configuration, wherein the SCR/MBS applies to all cells and not just the CLP=0 cells, is the same as the algorithm in Figure 7, except the second decision box (the second diamond) that tests "the cell has CLP=0?" is omitted. And in the last box, the "conforming cell" box, the text "If arriving cell had CLP=0 then" is also omitted. Likewise, the algorithm for the third configuration is the same as the algorithm in Figure 7, except that in the third decision box (the third diamond) the arrow corresponding to "No" does not go to the "non-conforming cell" box but rather goes to a box wherein the CLP bit is changed to 1, and from which an egress arrow goes to the "conforming cell" box

Figure 7. Generic Cell Rate Algorithm for Peak Cell Rate on aggregate CLP=0+1 cell flow and for Sustainable Cell Rate on CLP=0 cell flow. (From Ref. [8].)

The VBR service category can be partitioned into real-time and non-real-time. In the ATM Forum, the real-time VBR service category is for applications requiring constrained delay and delay variation, as would be appropriate for voice and video applications, [19]. Non-real-time VBR makes no commitment on delay but does include a commitment on cell loss. Non-real-time VBR service is analogous to Frame Relay service, which is popular for enterprise networks for connectivity between business locations.

As with CBR, key traffic controls used in the VBR service category include the negotiation of the PCR and SCR/MBS traffic parameters or the declaration of the Service Type and the subsequent policing of the submitted cell flow. In addition, the connection admission control (CAC) policy used in VBR enables the network operator to attain a statistical multiplexing gain, as compared with a CAC policy that allocates the PCR to the connections, while still meeting QoS commitments. Many such CAC policies are possible and the choice is at the discretion of the network operator. A network operator could choose a conservative policy that is based on the "worst case" traffic allowed by the PCR and SCR/MBS traffic parameters of the new and the established connections. Less conservative policies could use measurements of traffic of currently established connections and/or historical measurements of previous VBR connections. The historical measurements could be used to determine and/or validate stochastic models of the source traffic, and these models in turn could be used by the CAC policy. A popular concept is effective bandwidth where the variable-rate connection is viewed to have a constant rate, its effective bandwidth, that in some sense captures the stress the connection places on network resources. For example work on CAC see [26-29].

4.6.3 Available Bit Rate

The Available Bit Rate (ABR) service category is designed for applications that can adapt their information transfer rate based on feedback information from the network; the feedback information is conveyed in a special control cell, called the Resource Management (RM) cell. The RM cell can be distinguished from user cells via the Payload Type field in the ATM cell header, see Table 2 above. In the ABR service category, the RM cell provides the source with an indication of the currently available bandwidth for the connection. Closed-loop traffic controls are a basic feature of this service category.

In ABR, a source will specify a PCR and a minimum cell rate, where the latter may be zero. The bandwidth available for the connection may become as small as the minimum cell rate. The network commits to a QoS specified on the cell loss ratio, given that the cell flow is conforming at public interfaces, which would be the case if the source and destination follow a specified reference behavior in response to receiving a RM cell. The network also makes the relative assurance that for connections that share the same path, no connection shall be arbitrarily discriminated against nor favored. The Generic Cell Rate Algorithm (GCRA) is again used to determined conformance, though in a dynamic form where the increment parameter, I, varies over time to follow the currently allowed cell rate on the connection, see [19] or [30] for details.

As a rough summary of the ABR control procedures, a source determines the current bandwidth available for the connection by periodically emitting RM cells. Some of the fields of the RM cell, in particular the congestion-indication bit, the no-increase bit, and/or the explicit-rate field, are modified at network nodes to indicate to the source a revised allowed rate. When the destination receives the RM cell, the destination sends the RM cell back to the source (on the companion connection in the reverse direction). When the source receives the RM cell, the source updates the allowed sending rate, based on the information in the RM cell. There are many more aspects to the control; see [19] and [30] for the complete specification, and see K. Fendick [31] for a more readable summary.

4.6.4 ATM Block Transfer

The ATM Block Transfer (ABT) service category introduces the concept of a block of cells that is delineated by RM cells. The ABT service category transports complete blocks with low cell loss and low cell delay variation, comparable to that from the CBR service category. In a typical case, a higher-layer protocol data unit would be packaged as an ATM block, though an ATM block may also contain multiple higher-layer data units. In the ABT service category bandwidth is allocated on a block by block basis. Two service categories within ABT are defined. In ABT with Delayed Transmission the source sends an RM cell to request a rate at which to transmit cells of a block, called the Block Cell Rate, and then the source waits for a response RM cell from the network before sending the block of user data. In ABT with Immediate Transmission (ABT/IT), a user wishing to transmit an ATM block sends an RM cell followed by the user-data cells. If a network node along the connection can not support the requested rate, the request is denied, and in the case of ABT/IT the cells of the block may be discarded. At connection establishment the user may negotiate a sustainable cell rate traffic descriptor to obtain a guaranteed bandwidth for the connection. In the ABT service category, the SCR is at the block level, as opposed to the cell level as in VBR. Again, the Generic Cell Rate Algorithm (GCRA) is used to define conformance at public interfaces, where the increment parameter, I, varies over time to track the current Block Cell Rate. Again, there are many more aspects to ABT, and for a complete specification see [8] and [30].

In the ABT framework, resource allocation within the network is block oriented - resources needed for the transfer of an ATM block are dynamically allocated on a block basis. The network operator engineers resources and implements a connection admission control policy to keep the probability a request is denied within a specified level. At this time, ABT has not been generally implemented.

4.6.5 Unspecified Bit Rate

The Unspecified Bit Rate (UBR) service category is intended for delay-tolerant applications. In contrast to ABR, UBR does not use a feedback traffic-control mechanism. However, such mechanisms may be operating at a layer above the ATM layer, such as the Transmission Control Protocol (TCP) running on top of the Internet Protocol (IP), which in turn is running over a UBR connection. Although the network operator may engineer resources to support UBR connections, the specification of UBR does not include QoS commitments on cell loss or cell delay. UBR can be viewed as a simple cell relay service analogous to the common term “best effort service.” The PCR traffic parameter is negotiated for a UBR connection as it may identify a physical bandwidth limitation of the application or of a link along the path of the connections. However, the PCR is not necessarily policed by the public network. Network nodes that support UBR connections would need to isolate in some fashion the non-UBR connections (if present) from the UBR connections. In addition, a desirable feature would be some isolation of each UBR connection from the other UBR connections. This could be accomplished via weighted fair queueing and buffer management schemes.

4.6.6 Summary of ATM Service Categories

The following table is based on Table 2-1 in [19] and expanded to include the ABT service category specified in [8].

Table 5. Summary of ATM Service Categories.[*]

4.7 Suitability of ATM Service Categories for Transporting Video

Any of the above five service categories could be used to support the transport of video – some more effectively than others.

To begin with the least likely candidate, although the Unspecified Bit Rate (UBR) service category provides no loss or delay guarantees, if the network is lightly loaded then the loss and delay will indeed be low. Just as Ethernet LAN's can support some types of video, as LANs are typically lightly loaded, so too could a UBR VCC, in a similarly lightly loaded network. Of course, if many video transmissions are attempted, then the assumption of light load would be violated. A UBR VCC, as does any VCC, has the desirable attribute of delivering the transmitted bits in order. Also, a UBR VCC places few requirements on the video system – no need to shape the encoded bits to a traffic descriptor, or to react to control messages (resource management cells) from the network.

---

[*] Copyright 1996 The ATM Forum.

In contrast, a CBR VCC is similar to a circuit-switched connection, or private line, in traditional public telephone networks. Thus, a CBR VCC is a natural choice for transporting constant-rate encoded video, where the Peak Cell Rate of the CBR VCC is chosen to match the bit rate of the encoded video. Besides duplicating traditional private-line service, a CBR VCC has the additional flexibility that the requested rate can be whatever is appropriate for the source, and is no longer constrained to the rather coarse granularity of traditional services. Also, a CBR VCC can be used to transport variable-rate encoded video, given that the maximum bit-rate of the video is no more than the capacity of the CBR VCC. Moreover, the maximum bit-rate can be reduced via smoothing of the encoded video over a few frame times. The smoothing could be done in the encoder's buffer, or at the ATM level at the beginning of the ATM connection, or, as a service to the user, at the ingress to a service provider's network. However, using a CBR VCC for variable-rate video leads to a relatively inefficient use of the bandwidth. The remaining three service categories offer solutions for transporting variable-rate video while obtaining greater efficiency of network resources, as compared with using a CBR VCC.

The real-time VBR service category was defined to support voice and video applications. The specifics of how video transport can effectively use real-time VBR connections have attracted a good deal of research interest. For a recent review see Lakshman, Ortega, and Reibman [32]. A simple, natural idea is for the encoder to hold the quantizer step-size constant, to eliminate the encoder buffer and to transmit the encoded video as soon as possible to reduce delay in the end-system. Unfortunately, it is difficult to select values a priori for Sustainable Cell Rate and Maximum Burst Size such that the flow of emitted cells is indeed conforming. If the video is pre-recorded, as in a play-back application, then an off-line computation can be done to determine the appropriate traffic parameters. There is a continuum of choices. For each value of the SCR that is between the average and peak bit rate of the video, there is a minimum value for the MBS for which the video transmission would be conforming. However these parameters are fairly "loose" in the sense that if the network operator were to use them for connection admission control, the network resources would be inefficiently used, leading, in principle, to higher prices for the transport of the video. Some improvement is obtained if the encoded video, prior to transmission, is smoothed in some fashion over a time scale of a few frame times, see [33] for an illustrative study.

Video and real-time VBR connections are better matched to one another if the video encoder knows the traffic descriptor of the ATM connection and suitably adapts the encoding. Heuristically, one can think of the leaky-bucket algorithm for the SCR/MBS traffic parameters as a virtual extension to the physical encoder buffer. The video terminal emits cells and keeps track of the resulting state of the lucky-bucket, (which corresponds to the content in a virtual buffer, "virtual" because the cells are not actually being buffered). When the state of the lucky-bucket is close to being filled, the video terminal then begins to physically buffer encoded bits, and thus allows the state of the leaky-bucket to decrease. When the physical buffer is getting full, the encoder can adjust the encoding such as by increasing the quantizer step size. Various authors have proposed particular schemes that have additional desirable properties such as controlling the end-to-end delay (including the delay in the end-systems), maintaining the quality of the video, and creating traffic streams that enable efficient use of network resources. For recent examples, see Hsu, Ortega and Reibman [34], and Hamid, Roberts and Rolin [35], as well as references in [32].

The ATM Block Transfer (ABT) service category can also be used for the transport of variable-rate video. The basic idea would be that the source would request a Block Cell Rate that is currently needed, say on a scene by scene basis, or possibly a frame by frame basis. Of course, the source would need to determine what should be the requested rate. This is relatively straight forward if the encoding is already completed, such as with stored (play-back) video. The rate could also be determined in real-time based on the current buffer occupancy. If ABT with Delayed Transmission is used, then the source needs to wait for an acknowledgement before transporting (if the request is for a higher rate), which might be preferable to ABT with Immediate Transmission, where transmitted cells could be discarded if the requested rate is not granted. To have an assurance that a requested rate would be granted, the ABT connection should be established with the block-level SCR. Given that the source stays within the SCR, then requested rate changes, in particular requests to increase the rate, would be granted. Of course, there is the issue of determining the appropriate choice for the SCR.

The ABR service category was originally defined for non-real-time data applications. However, the feedback information on the rate that the network can currently support, combined with a minimum cell rate can be used by a video encoder to good effect. For illustrative schemes on this theme see Lakshman, Mishra, and Ramakrishnan [36] and Duffield, Ramakrishnan, and Reibman [37]. And in particular see Chapter 10 of the present volume which pursues this theme in detail.

4.8 Where is ATM?

Regarding Figure 2, we mentioned that the end-point of the ATM connection may or may not be at the user's end-system. Currently ATM is mainly deployed in the backbone network of service providers, as well as within large interconnection points, called network access points, between internet service providers. Even when the ATM network is within the backbone, and does not extend to the end-systems, ATM can still be used to transport video. A given ATM connection could be used to transport the aggregate video traffic between two backbone nodes. The ATM connection might reasonably be a CBR or a real-time VBR connection. Multiple video streams could be put on one VCC, or each video stream could be put on its own VCC, and the aggregate carried as a VPC across the backbone. The ATM network doesn't need to know that the user-data is actually encoded video, but only that the given packets are to be transported on ATM connections with specified QoS objectives particularly with regard to delay variation and loss.

If the ATM connection did extend to the end-system, then the video encoding and the ATM connection can be tailored for one another. For real-time VBR connections, the various shaping, buffering, and encoding procedures referenced in Section 4.7 could be used. For ABR connections, the video encoder could also exploit feedback information from the network.

ATM could begin to become popular in end-systems in conjunction with higher speed access to the internet from the home and business. One of the contending technologies for internet access is Asymmetric Digital Subscriber Line (ADSL), which uses the current telephone wires and operates at speeds up to a few Megabits per second in the direction towards the end-system and at speeds of some hundreds of Kilobits per second in the direction from the end-system. The asymmetry in the bandwidths takes advantage of the asymmetry in traffic load for typical user applications of web browsing. One variant of ADSL is being developed by the Universal ADSL Working Group and is called Universal ADSL (U-ADSL). U-ADSL is designed for easy and early deployment, and supports speeds up to 1.5Mbps to the home and 512Kbps from the home. U-ADSL has been standardized in ITU Recommendation G.992.2, [38], and in which the support of ATM is a requirement. On the software front, Microsoft has announced that its Windows 2000 Professional software and the next major release of NT Server will support ATM. Sun Microsystems already provides ATM network interface cards running at speeds of 155Mbps and

622Mbps. Some telephone companies have begun to offer ADSL service, and some of the major personal computer manufacturers have ADSL modems available. A detailed discussion of running ATM over ADSL is provided by Kwok in [39].

4.9 Further Reading

This chapter has provided a basic introduction to ATM networks. References were given to source material, which mainly consists of ITU and ATM Forum documents. Although this material constitutes the primary source, it was not written to be explanatory or educational, and thus is not particularly readable for someone who is not already knowledgeable in ATM (and sometimes even prior expertise doesn't help). If the reader is inclined to learn more about ATM, there is a wealth of secondary source material in the form of readable textbooks. The following is a somewhat arbitrary sampling – all provide a good, overall introduction and summary of ATM. A recent book, published in 1999, is by Ginsberg, [40], who focuses on internetworking and services, and includes a discussion of MPLS. Black has written a three-volume series on ATM, the first covering the foundations of broadband networks, [41], the second on signaling, [42], and the third on internetworking, [43]. Handel, Huber, and Schroder, 1998, provide a thorough and readable treatment, [44]. Onvural, 1995, also provides thorough coverage and has a focus on performance issues, [5]. Dyson and Spohn, 1995, include a discussion of ATM hardware, software and end-systems, [45]. A classic text, revised in 1995, is by Prycker, [46]. If the reader has a particular interest in topics of traffic characterization, connection admission policies, and network design, an excellent text is by Roberts, Mocci, Virtamo, editors, [26]. Kwok focuses on ATM in access networks to the home and business, including multicasting of video, [33].

4.10 Acronyms Used in the Chapter

AAL	ATM Adaptation Layer
ABR	Available Bit Rate
ABT	ATM Block Transfer
ADSL	Asymmetrical Digital Subscriber Line
ATM	Asynchronous Transfer Mode
B-ISDN	Broadband Integrated Services Digital Network
BUS	Broadcast and Unknown Server
CAC	Connection Admission Control
CBR	Constant Bit Rate
CDV	Cell Delay Variation
CDVT	Cell Delay Variation Tolerance
CLR	Cell Loss Ratio
CPCS	Common Part of the Convergence Sublayer
CTD	Cell Transfer Delay
DBR	Deterministic Bit Rate
DSL	Digital Subscriber Line
GCRA	Generic Cell Rate Algorithm
GFR	Guaranteed Frame Rate

HEC	Header Error Control
IP	Internet Protocol
ITU	International Telecommunications Union
LAN	Local Area Network
maxCTD	maximum Cell Transfer Delay
Mbps	Mega (10^6) bits per second
MBS	Maximum Burst Size
MPLS	Multi-Protocol Label Switching
MPOA	Multi-Protocol Over ATM
NNI	Network-Network Interface, or Network-Node Interface
OAM	Operations Administration and Maintenance
PCR	Peak Cell Rate
PDU	Protocol Data Unit
PHY	Physical layer
PT	Payload Type
QoS	Quality of Service
RM	Resource Management
SAR	Segmentation and Reassembly
SBR	Statistical Bit Rate

SCR	Sustainable Cell Rate
SDH	Synchronous Digital Hierarchy
SONET	Synchronous Optical NETwork
TCP	Transmission Control Protocol
TD	Traffic Descriptor
U-ADSL	Universal – Asymmetric Digital Subscriber Line
UBR	Unspecified Bit Rate
UNI	User-Network Interface
UTP	Unshielded Twisted Pair
VBR	Variable Bit Rate
VCC	Virtual Channel Connection
VCI	Virtual Channel Identifier
VPC	Virtual Path Connection
VPI	Virtual Path Identifier

References

[1] ITU-T Recommendation Q.2931, Broadband Integrated Services Digital Network (B-ISDN) - Digital subscriber signalling system no. 2 (DSS 2) - User-network interface (UNI) - Layer 3 specification for basic call/connection control, Inter. Telecommunications Union, Geneva Switzerland, February, 1995.

[2] ATM Forum ATM User-Network Interface (UNI) Signalling Specification Version 4.0, ATM Forum Technical Committee, July, 1996

[3] ATM Forum Private Network-Network Interface Specification Version 1.0 (PNNI 1.0), ATM Forum Technical Committee, April, 1996.

[4] ITU-T Recommendation I.432, B-ISDN User Network Interface Specification, Inter. Telecommunications Union, Geneva Switzerland, August, 1996

[5] R Onvural. Asynchronous Transfer Mode Networks, Performance Issues. Norwood, Mass. U.S.A.: Artech House, 1995.

[6] Recommendation I.361, B-ISDN ATM Layer Specification, Inter. Telecommunications Union, Geneva, Switzerland, November, 1995.

[7] ITU-T Recommendation I.610, B-ISDN Operations and Maintenance Principles and Functions, Inter. Telecommunications Union, Geneva, Switzerland, November 1995.

[8] ITU-T Recommendation I.371, Traffic Control and Congestion Control in B-ISDN, Inter. Telecommunications Union, Geneva, Switzerland, May, 1996.

[9] ITU-T Recommendation E.735, Framework for Traffic Control and Dimensioning in B-ISDN, Inter. Telecommunications Union, Geneva, Switzerland, May, 1997.

[10] ATM Forum, LAN Emulation Over ATM Specification, Version 1.0, ATM Forum Specification, January, 1995.

[11] ATM Forum, Multi-Protocol Over ATM Specification Version 1.0, ATM Forum Specification, July, 1997.

[12] ITU-T Recommendation I.363.1, B-ISDN ATM Adaptation Layer Specification: Type 1 AAL, Inter. Telecommunications Union, Geneva, Switzerland, August 1996.

[13] ITU-T Recommendation I.363.2, B-ISDN ATM Adaptation Layer Specification: Type 2 AAL, Inter. Telecommunications Union, Geneva, Switzerland, September, 1997.

[14] JH Baldwin, BH Bharucha, BT Doshi, S Dravida, S Nanda. A new ATM adaptation layer for small packet encapsulation. Bell Labs Technical Journal, 2:111-131,1997.

[15] K Sriram, YT Wang. Voice over ATM using AAL2 and bit dropping: performance and call admission control. IEEE J. on Selected Areas on Communications 17:18-28, 1999.

[16] ITU-T Recommendation I.363.5, B-ISDN ATM Adaptation Layer Specification: Type 5 AAL, Inter. Telecommunications Union, Geneva, Switzerland, August, 1996.

[17] ITU-T Recommendation I.358, Call Processing Performance for Switched Virtual Channel Connections in a B-ISDN, Inter. Telecommunications Union, Geneva, Switzerland, June, 1998.

[18] ITU-T Recommendation I.356, B-ISDN ATM Layer Cell Transfer Performance, Inter. Telecommunications Union, Geneva, Switzerland, May, 1996.

[19] ATM Forum Traffic Management Specification Version 4.0, ATM Forum Technical Committee, April, 1996.

[20] H Heiss. QoS Results for the GFR ATM Service. In: P Key, D. Smith, eds. Teletraffic Engineering in a Competitive World, Proceedings of the 16th Inter. Teletraffic Congress. Amsterdam: Elsevier, 1999, pp 227-236.

[21] R Guerin J Heinanen. UBR+ Service Category Definition. ATM Forum contribution 96-1598, December 1996. UBR+ Enhancements. ATM Forum contribution 97-0015, February 1997.

[22] AW Berger, W Whitt. The Brownian Approximation for Rate-Control Throttles and the G/G/1/C Queue. Discrete Event Dynamic Systems 2:7-60, 1992.

[23] AW Berger, W Whitt. A Comparison of the Sliding Window and the Leaky Bucket. Queueing Systems 20:117-138, 1995.

[24] AW Berger. Desirable Properties of Traffic Descriptors for ATM Connections in a Broadband ISDN. In: J Labetoulle, JW Roberts, ed. The Fundamental Role of Teletraffic in the Evolution of Telecommunications Networks, Proceedings of the 14th Inter. Teletraffic Congress. Amsterdam: Elsevier, 1994, pp 233-242.

[25] AW Berger, AE Eckberg. A B-ISDN/ATM Traffic Descriptor and Its Use in Traffic and Congestion Controls. Proceedings of GLOBECOM, Phoenix, Arizona, 1991, pp 266-270.

[26] J Roberts, U Mocci, J. Virtamo eds. Broadband Network Traffic, Performance Evaluation and Design of Multiservice Networks. Berlin: Springer-Verlag, 1996.

[27] FP Kelly. Notes on Effective Bandwidths. In: FP Kelly, S Zachary, I. Ziedins eds. Stochastic Networks. Oxford: Claredon Press, 1996, pp 141-168.

[28] EW Knightly, NB Shroff. Admission Control for Statistical QoS: Theory and Practice. IEEE Network 13(2):20-29, March 1999.

[29] AW Berger, W Whitt. Extending the Effective Bandwidth Concept to Networks with Priority Classes. IEEE Communications Magazine 36:78-83, 1998.

[30] ITU-T Recommendation I.371.1, Traffic Control and Congestion Control in B-ISDN, Conformance Definitions for ABT and ABR, Inter. Telecommunications Union, Geneva, Switzerland, June, 1997.

[31] K Fendick. Evolution of Controls for the Available Bit Rate Service. IEEE Communications Magazine 34:35-39, 1996.

[32] TV Lakshman, A Ortega, AR Reibman. VBR video: tradeoffs and potentials. Proceedings of the IEEE 86:952-973, 1998.

[33] AR Reibman, AW Berger. Traffic Descriptors for VBR Video Teleconferencing over ATM Networks. IEEE/ACM Transactions on Networking 3:329-339, 1995.

[34] CY Hsu, A Ortega, AR Reibman. Joint Selection of Source and Channel Rate for VBR Video Transmission Under ATM Policing Constraints. IEEE J. on Selected Areas on Communications 15:1016-1028, 1997.

[35] M Hamdi, JW Roberts, P Rolin. Rate Control for VBR Video Coders and Broadband Networks. IEEE J. on Selected Areas on Communications 15:1040-1051, 1997.

[36] TV Lakshman, PP Mishra, KK Ramakrishnan. Transporting compressed video over ATM networks with explicit rate feedback control. Proceedings IEEE Infocom 1997, Kobe Japan, 1997, pp 38-47.

[37] NG Duffield, KK Ramakrishnan, AR Reibman. SAVE: An Algorithm For Smoothed Adaptive Video Over Explicit Rate Networks. IEEE/ACM Transactions on Networking 6:717-728, 1998.

[38] ITU-T Recommendation G.992.2, Splitterless Asymmetrical Digital Subscriber Line (ADSL) Transmissions, Inter. Telecommunications Union, Geneva, Switzerland, COM 15-R 27-E determined for approval October, 1998.

[39] T Kwok. ATM: The New Paradigm for Internet, Intranet and Residential Broadband Services and Applications. Prentice Hall, 1998.

[40] D Ginsberg. ATM Solutions for Enterprise Internetworking. Harlow, U.K.: Addison-Wesley, 1999.

[41] UD Black. ATM: Foundations for Broadband Networks, Vol. I. Englewood Cliffs, NJ: Prentice Hall, 1996.

[42] UD Black. ATM: Signaling in Broadband Networks, Vol. II. Englewood Cliffs, NJ: Prentice Hall, 1997.

[43] UD Black. ATM: Internetworking, Vol. III. Englewood Cliffs, NJ: Prentice Hall, 1997.

[44] R Handel, MN Huber, S Schroder. ATM Networks Concepts Protocols Applications. Harlow, U.K.: Addison-Wesley, 1998.

[45] DE McDyson, DL Spohn. ATM Theory and Application. New York, NY: McGraw-Hill, 1995.

[46] MD Prycker. Asynchronous Transfer Mode: Solution for Broadband ISDN. Hempstead U.K.: Prentice Hall, 1995.

List of Captions to Tables

Table 1. Example physical interfaces for which an ATM Physical Layer is defined.

Table 2. Standardized code points in the Payload Type field.

Table 3. Provisional Performance Objectives in ITU Recommendation I.356.

Table 4. Comparison of vocabulary in the ITU and ATM.

Table 5. Summary of ATM Service Categories.

List of Captions to Figures

Figure 1. ATM Cell

Figure 2. Protocol stack in ATM networks

Figure 3. ATM cell header at the User-Network Interface (UNI)

Figure 4. Transmission Path supports multiple VPCs, which in turn support multiple VCCs.

Figure 5. Illustration of label swapping of VPI/VCI values for a given Virtual Channel Connection.

Figure 6. Generic Cell Rate Algorithm (From Ref. [8])

Figure 7. Generic Cell Rate Algorithm for Peak Cell Rate on aggregate CLP=0+1 cell flow and for Sustainable Cell Rate on CLP=0 cell flow. (From Ref. [8].)

Physical Layer Interface	Transmission Rate [Mbps]	Medium
DS-1 (T-1)	1.544	Coax
E-1	2.048	Coax
DS-3 (T-3)	44.736	Coax
E-3	34.368	Coax
SDH STM-1, SONET OC-3c	155.52	Fiber
SDH STM-4c, SONET OC-12c	622.08	Fiber
SDH STM-16c, SONET OC-48c	2,488.32	Fiber
FDDI	100.	Fiber
Raw cells	155.52	Fiber
Raw cells	622.08	Fiber
Raw cells	51.84	Unshielded Twist Pair copper wire, UTP-3

Table 1.

Payload type coding	Interpretation
000	User data cell, congestion not experienced. ATM user-to-ATM-user indication = 0
001	User data cell, congestion not experienced. ATM user-to-ATM-user indication = 1
010	User data cell, congestion experienced. ATM user-to-ATM-user indication = 0
011	User data cell, congestion experienced. ATM user-to-ATM-user indication = 1
100	OAM F5 segment associated cell
101	OAM F5 end-to-end associated cell
110	Resource management cell
111	Reserved for future VC functions

Table 2.

QoS Class	ITU Rec. I.356 Provisional Performance Objectives:			
	upper bound on the mean cell transfer delay	upper bound on the difference between the upper and lower 10^{-8} quanitles of cell transfer delay	upper bound on the cell loss probability (regardless of the value of the CLP bit in the cell header)	upper bound on the cell loss probability for cells with CLP bit = 0
Stringent class	400 msec	3 msec	$3*10^{-7}$	none
Tolerant class	U	U	10^{-5}	none
Bi-level class	U	U	U	10^{-5}
"U" class	U	U	U	U

Table 3.

Inter. Telecommunications Union	ATM Forum
ATM transfer capability	Service category
Deterministic Bit Rate	Constant Bit Rate
Statistical Bit Rate	Variable Bit Rate
Available Bit Rate	Available Bit Rate
ATM Block Transfer	-
-	Unspecified Bit Rate
ATM Performance Parameters	Quality of Service Parameters

Table 4..

	ATM Service Category:					
Attribute	CBR	VBR real-time	VBR non-real-time	ABR	UBR	ABT
Cell loss ratio	specified, note 1			specified note 2	unspecified	specified note 3
Cell Transfer Delay and Delay Variation	specified		un-specified	unspecified note 5	unspecified	specified
Peak Cell Rate	specified			specified note 4	specified note 5	specified
SCR/MBS	not applicable	specified		not applicable		specified note 6
Real-time Control via RM cells	no			yes	no	yes

Note 1	For CBR and VBR, the cell loss ratio may be unspecified for CLP=1 cells.
Note 2	Cell loss ratio is minimized for sources that adjust their cell flow in response to control information.
Note 3	For conforming blocks, the cell loss ratio is comparable to that for CBR. Block loss ratio is specified.
Note 4	Represents the maximum rate at which the source may ever send. The momentary maximum allowed rate is subject to control information.
Note 5	Peak Cell Rate is not subject to CAC and UPC procedures.
Note 6	SCR specified for ATM Blocks.

Table 5.

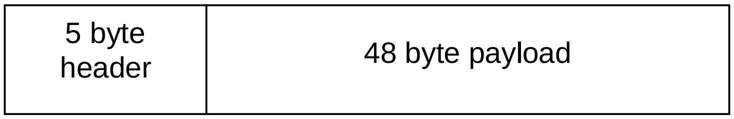


Figure 1

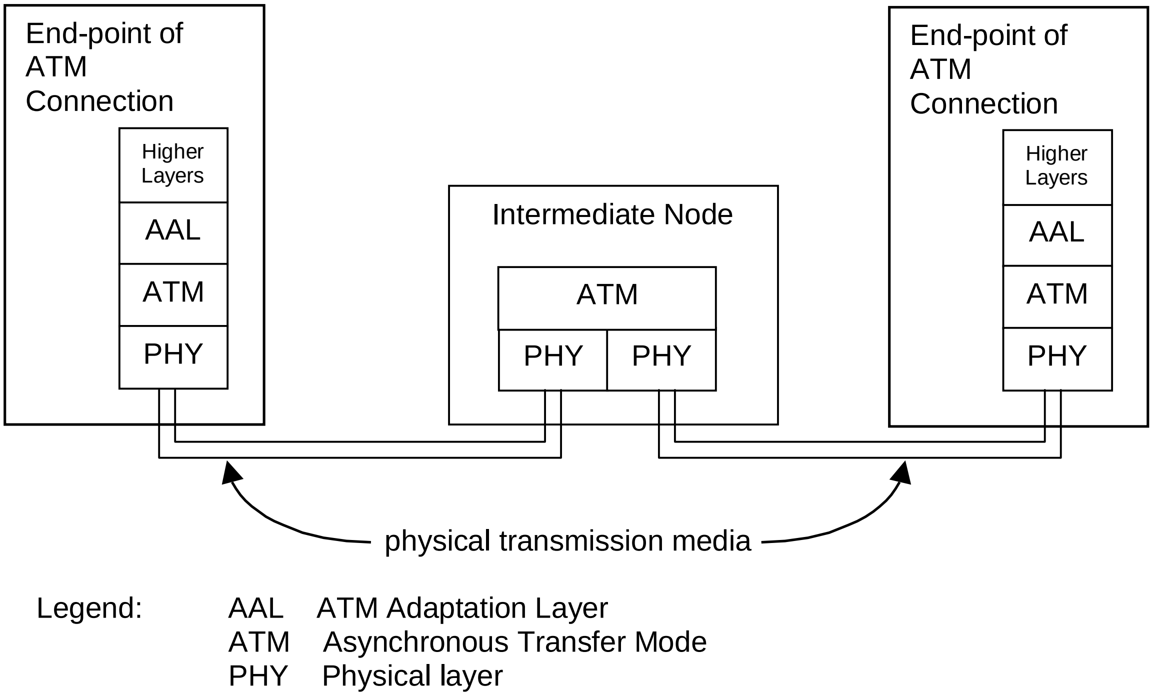


Figure 2

8	7	6	5	4	3	2	1	Bit Octet
GFC				VPI				1
PI				VCI				2
VCI								3
VCI				PT			CLP	4
HEC								5

Figure 3

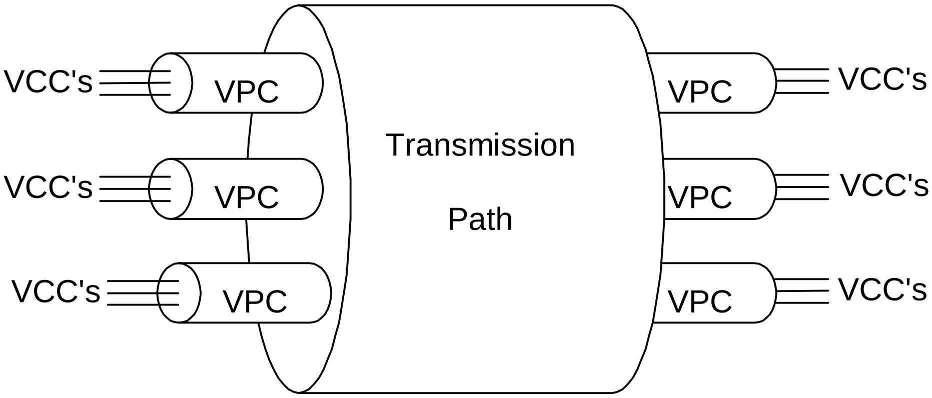
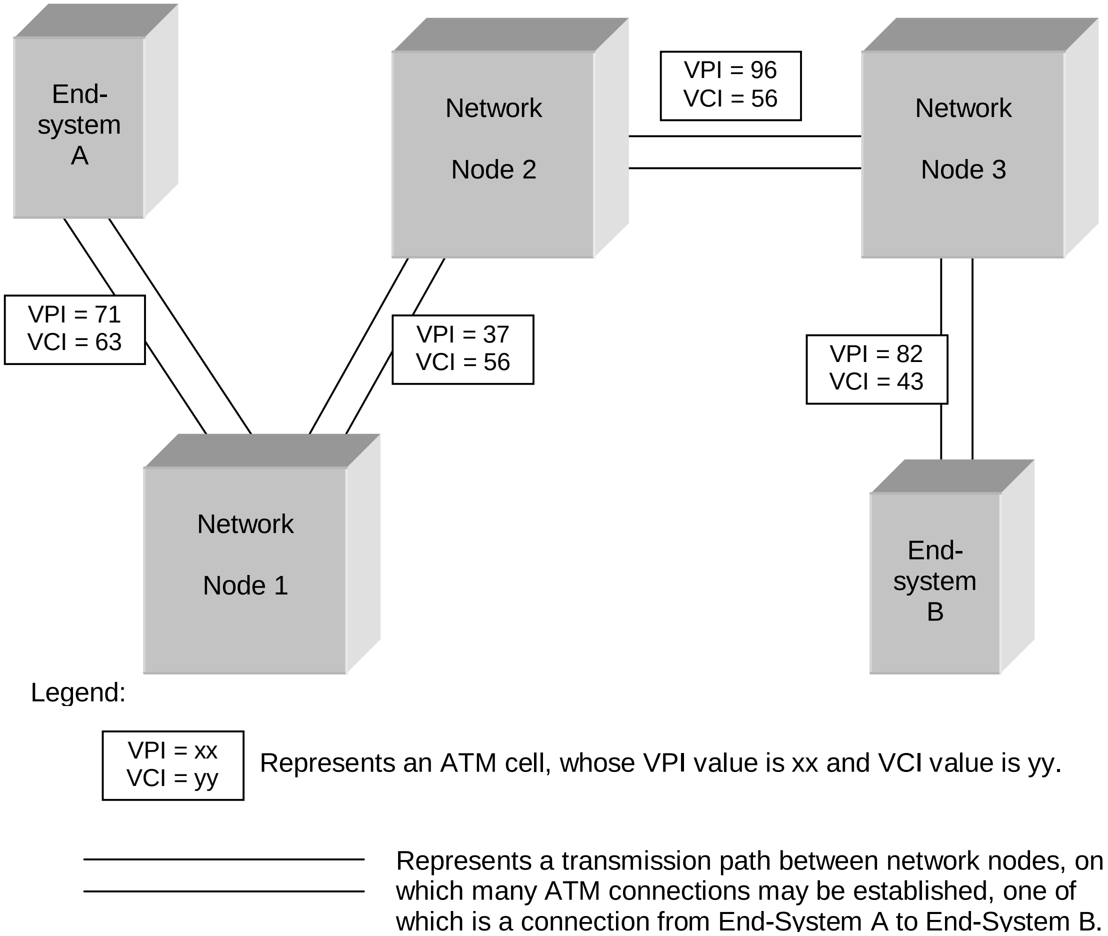

VCC's
VPC
VCC's
VPC
VCC's
VPC
Transmission
Path
VPC
VCC's
VPC
VCC's
VPC
VCC's

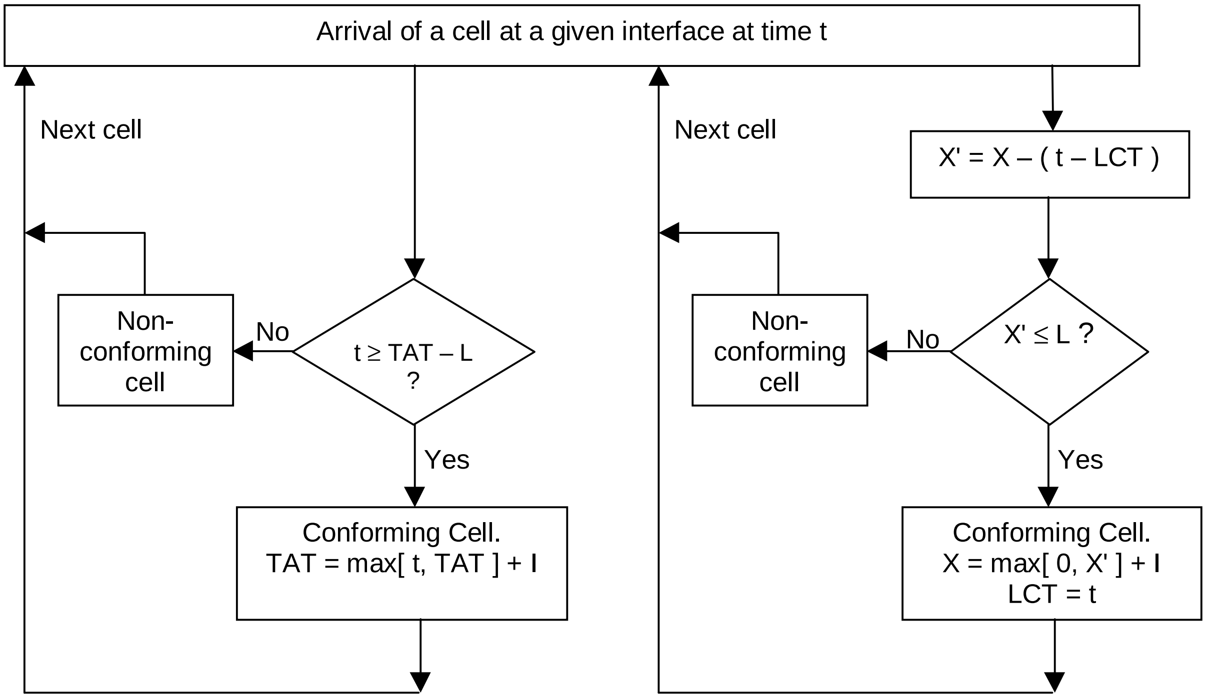
Figure 4.



Represents an ATM cell, whose VPI value is xx and VCI value is yy.

Represents a transmission path between network nodes, on which many ATM connections may be established, one of which is a connection from End-System A to End-System B.

Figure 5.



t	Time of arrival of cell
I	Reciprocal of peak cell rate
L	Cell Delay Variation tolerance

Virtual scheduling algorithm

TAT	Theoretical Arrival Time for PCR traffic descriptor (TD)

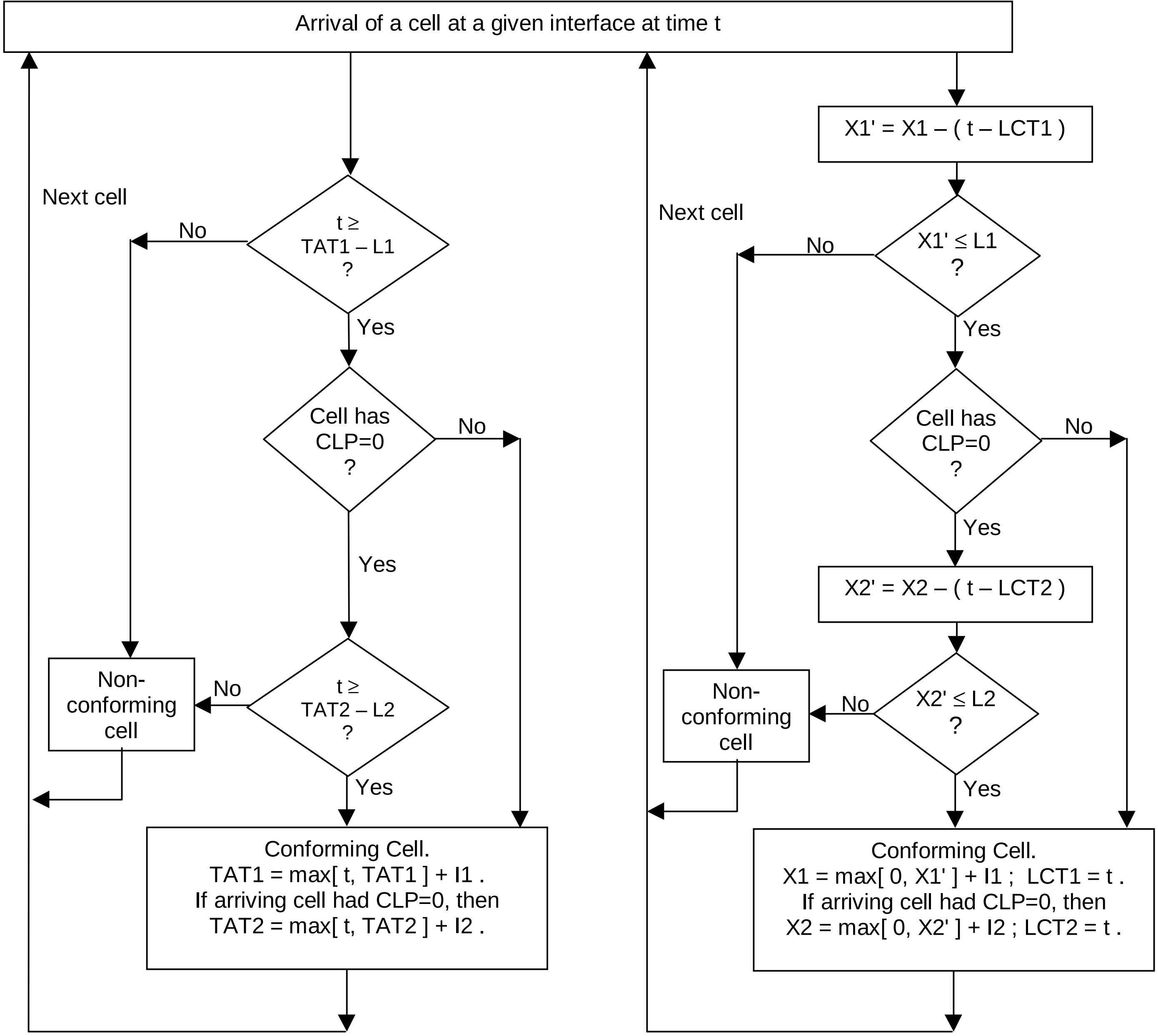
At the arrival time of the first cell of the connection, TAT = t.

Continuous-State Leaky bucket algorithm

X	Value of leaky bucket content for PCR traffic descriptor (TD)
LCT	Last Conformance Time for PCR TD
X'	Auxiliary variable

At the arrival time of the first cell of the connection, X = 0 and LCT = t.

Figure 6



t	Time of arrival of cell
I1, I2	Reciprocal of peak cell rate and sustainable cell rate, respectively
L1, L2	Cell Delay Variation tolerance and Intrinsic Burst Tolerance

Virtual scheduling algorithm

TAT1	Theoretical Arrival Time for PCR traffic descriptor (TD)
TAT2	Theoretical Arrival Time for SCR traffic descriptor (TD)

At the arrival time of the first cell of the connection, TAT1 = TAT2 = t.

Continuous-State Leaky bucket algorithm

X1, X2	Value of leaky bucket content for PCR and SCR TD's respectively
LCT1, LCT2	Last Conformance Times for PCR and SCR TD's respectively
X1', X2'	Auxiliary variables

At the arrival time of the first cell of the connection, X1 = X2 = 0 and LCT1 = LCT2 = t.

Figure 7