MPEG-2 Transport over ATM Networks

Christos Tryfonas
Anujan Varma

UCSC-CRL-99-01
February 2, 1999

Computer Engineering Department
University of California, Santa Cruz
Santa Cruz, CA 95064

ABSTRACT

The growth of Asynchronous Transfer Mode (ATM) technology and the maturity of coding standards have made the large-scale deployment of high-quality audiovisual services possible. In this article, we discuss the principal issues involved in transporting MPEG-2 streams over ATM networks for both constant bit-rate (CBR) and variable bit-rate (VBR) MPEG-2 Transport Streams. We review existing and proposed approaches to deal with these issues and evaluate some of them by simulation to demonstrate the important tradeoffs in the design of networked audiovisual systems.

Keywords: MPEG-2, ATM networks, set-top box, MPEG-2 Systems layer.

1 Introduction

The explosion of the Internet has created demand for new applications traditionally carried over circuit-switched networks. Such applications include audio telephony, video conferencing and video-on-demand (VoD) services. New standards are emerging to support these applications in the context of both connectionless and connection-oriented packet-switched networks.

Asynchronous Transfer Mode (ATM) is an emerging standard for broadband networks that allows a wide range of traffic types —ranging from real-time video to best-effort data— to be multiplexed in a single physical connection-oriented network. A key benefit of ATM technology is its ability to provide quality-of-service (QoS) guarantees to applications. These QoS guarantees are in the form of bounds on end-to-end delay, packet delay variation (jitter) and packet loss rate. Several classes of service have been defined in the context of ATM networks to satisfy the QoS needs of various applications. The Constant Bit-Rate (CBR) and Real-Time Variable Bit-Rate (RT-VBR) service classes are intended for real-time applications with stringent requirements on delay, jitter, and loss rate, such as video-on-demand (VoD) services. The Non-Real-Time Variable Bit-Rate (NRT-VBR) service class is intended for applications where no jitter control is needed, but a delay guarantee is still required. The Available Bit-Rate (ABR) service class is intended for delay-tolerant best-effort applications and uses a rate-based feedback approach to control potential congestion. Regular TCP applications fall under this service class. Finally, the Unspecified Bit-Rate service (UBR) does not offer any service guarantees and thus, has the lowest priority among all the classes. Within a service class, the feasibility of supporting a specified set of QoS requirements is determined by admission-control algorithms. Because of its ability to support a prespecified set of requirements, ATM technology is inherently well-suited for the design of Video Dial Tone (VDT) networks.

Besides the underlying network infrastructure, the coding method used for digital video and audio for VDT services has a significance influence on the viability and performance of such services. MPEG-2 is currently the most popular standard for audio and video compression in VDT networks [18, 19]. Being capable of exploiting both spatial and temporal redundancies, it achieves compression ratios up to 200:1 and can encode a video or audio source to almost any level of quality. MPEG-2 standard offers two ways of multiplexing elementary audio, video or private streams to form a program: the *MPEG-2 Program Stream* and the *MPEG-2 Transport Stream* formats. Although the MPEG-2 Program Stream format is used in the Digital Versatile Disk (DVD) standard for playback in stand-alone environments, it is not resilient to errors and therefore not suitable for transmission over error-prone environments. The MPEG-2 Transport Stream format is the approach suggested for transporting MPEG-2 over noisy environments, such as a packet network. Using explicit timestamps (called Program Clock References or PCRs in MPEG-2 terminology) that are inserted periodically into the stream, MPEG-2 Transport Streams ensure synchronization and continuity, and provide ways to facilitate the clock recovery at the decoder end.

In this article, we discuss the key problems in transporting MPEG-2 over ATM networks and describe solutions that have been proposed in the literature. We start by describing a generic architecture for MPEG-2 over ATM in Section 2. In Section 3, we discuss the key issues involved in transporting MPEG-2 over ATM, and the approaches proposed in the literature to deal with them for both the CBR and the VBR cases. In Section 4, we present some experimental results on how those issues influence the quality of the VDT service. Finally, we conclude with a summary in Section 5.

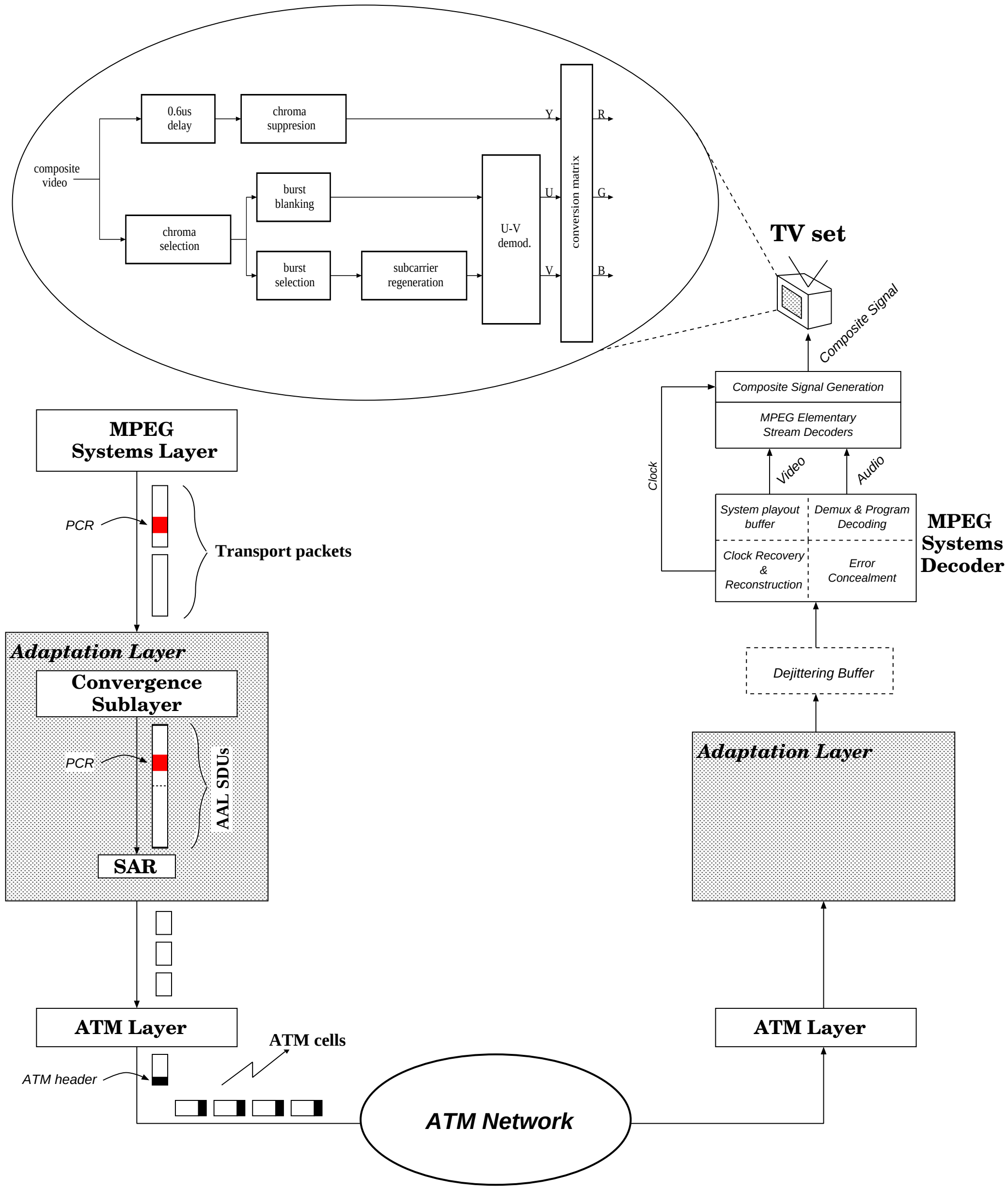

Figure 2.1: Generic architecture for MPEG-2 over ATM networks.

2 Protocol Architecture for MPEG-2 over ATM

A generic architecture for transporting MPEG-2 over ATM networks is illustrated in Figure 2.1. On the sender side, the MPEG-2 Transport Stream is sent to the network through an adaptation layer (AAL) and the ATM layer. On the receiver side, the architecture consists of the ATM layer, the adaptation layer, an optional dejittering buffer, the MPEG-2 Systems decoder, the MPEG-2 decoders for the elementary streams, and the TV set. The MPEG-2 Systems decoder includes among others, the phase-locked loop (PLL) used to recover the clock from the incoming PCR values for the synchronization of the sender and the receiver, and the system playout buffer.

The transport of MPEG-2 over ATM introduces several issues that must be addressed in order to deal with the problem on an end-to-end basis. These include the choice of the adaptation layer, method of encapsulation of MPEG-2 packets in AAL packets, service class selection in the ATM network for control of delay and jitter, and the design of the decoder. The choice of adaptation layer involves a number of tradeoffs [9]. The possible choices are Adaptation Layer 1 (AAL1), suitable for circuit-emulation type of services, Adaptation Layer 5 (AAL5), currently used for transporting data traffic with no real-time constraints, and Adaptation Layer 2 (AAL2) which is not standardized yet, but may offer an alternative in the future for transport of VBR MPEG-2 traffic. In the case of AAL5, two distinct approaches were proposed for encapsulation of MPEG-2 streams in AAL5 packets in the ATM Forum: the first approach is the *PCR-aware* scheme, in which the packetization is done ensuring that transport packets containing PCR values will be transmitted immediately. The second approach is the PCR-unaware scheme, where no distinction is made for packets containing PCR values; this may introduce significant jitter for PCR values during the encapsulation.

A wide range of proposals has been made for selecting the type of service under which MPEG-2 is to be transported over ATM [15, 23, 30, 52]. For constant bit-rate MPEG-2 streams, the CBR class of service is the natural choice. However, even in this case, the QoS provided by the ATM network may influence the overall quality significantly. For the variable bit-rate case, three main approaches have been proposed. The statistical service with rate renegotiation tries to maximize the multiplexing gain by capturing the VBR nature of MPEG-2 [15, 52]. According to this approach, the effective bandwidth of the source during a pre-determined interval is used to allocate resources in the network. If sufficient resources are not available the quality is degraded and in that sense, the service is statistical. This requires an algorithm to determine re-negotiation points. The second approach, based on a feedback-based Available Bit-Rate (ABR) service, uses feedback information to change the coding rate at the output of the MPEG-2 encoder to suit the available bandwidth [22, 23, 30]. In this approach, the service is considered best effort with some minimum guarantees. The last approach, which operates over a best-effort service such as that provided by the current Internet, the overall quality is almost entirely dependent on the congestion level of the network.

Synchronization issues may arise while transporting MPEG-2 over ATM due to cell delay variation (jitter). The presence of jitter introduced by the underlying ATM network may distort the reconstructed clock at the MPEG-2 audio/video decoder, which in turn may degrade the quality since the synchronization signals for display of the video frames are obtained from the recovered clock. A common solution is to use a dejittering mechanism at the receiver that absorbs any jitter introduced by the network.

In order to ensure acceptable quality at the receiver, each component of the end-to-end path must be designed to provide the desired level of service. Therefore, optimizing only specific components in the path may not be adequate for ensuring the desired quality for the viewer. For example, providing superior QoS in the ATM network may not be sufficient to maintain adequate quality at the receiver while using an adaptation layer that introduces significant jitter, and a poor phase-locked loop (PLL) design within the MPEG-2 decoder. Thus, the adaptation layer, encapsulation scheme, service class selection in the ATM network, dejittering mechanisms at the receiver and the PLL in the MPEG-2 system decoder must all be designed to provide the desired level of quality at the receiver. In the next section we proceed with a detailed discussion of several of these issues.

3 Issues in MPEG-2 Transport over ATM

The MPEG-2 standard [18] does not specify how an MPEG-2 Transport Stream is transported over a communication network. However, the timing model of MPEG-2 Systems Layer assumes a constant end-to-end delay from the encoder to the decoder end. This introduces a number of design issues that need to be addressed in order to ensure satisfactory quality at the receiver. Some of them are stated below:

1. Choice of Adaptation Layer.
2. Transport Packet Encapsulation.
3. Service Class Selection.
4. Clock Synchronization.

3.1 Choice of Adaptation Layer (AAL)

The Adaptation Layer is responsible for making the network behavior transparent to the application. It is divided into two sub-layers: the Segmentation and Reassembly (SAR) sublayer and the Convergence Sublayer (CS). The SAR sublayer is responsible for the segmentation of the outgoing Protocol Data Units (PDUs) into ATM cells and the reassembly of ATM cells back into the original PDUs. There are four types of adaptation layers currently defined for ATM networks: AAL1, AAL2, AAL3/4 and AAL5. Each of these is designed for supporting specific services and have different functionalities. The selection of a suitable adaptation layer for transporting MPEG-2 over ATM needs to take into account the specific requirements of MPEG-2 Transport Streams, such as jitter removal, error detection and/or correction, end-to-end delay minimization for real-time applications and support of both CBR and VBR applications.

Transport over AAL1

AAL1 was designed to support circuit emulation over ATM networks. It is ideally suited for transporting Constant Bit-Rate (CBR) traffic since it provides constant delay through the network using dejittering mechanisms at the destination. AAL1 provides two ways to synchronize the clocks and deliver a jitter-free clock at the receiver depending on the status of the CBR service clock. In the *synchronous* case, the service clock is assumed locked to a common network clock and its recovery is done directly from the network clock. In the *asynchronous* case, AAL1 provides two alternatives for recovering the clock at the receiver: the *Synchronous Residual Time Stamp* (SRTS) method and the *Adaptive Clock* method. In the former method, absolute clock information is exchanged for the synchronization whereas in the latter, the buffer fill levels are used in order to synchronize the transmitter and receiver. Besides, AAL1 offers the option of Forward Error Correction (FEC) that can hide the effects of cell losses in the network from the application.

AAL1 is the natural choice for transporting CBR MPEG-2 over ATM since the traffic has a constant bit-rate and needs constant end-to-end delay. However, there are a number of disadvantages with the use of AAL1 for MPEG-2 transport:

1. AAL1 cannot be used to carry Variable Bit-Rate (VBR) MPEG-2 Transport Streams which are likely to be dominant in the future.
2. The SRTS technique cannot be used if a common network clock is not available for use as the reference clock. Thus, AAL1 cannot be used in nationwide networks consisting of several carriers and unsynchronized clocks.

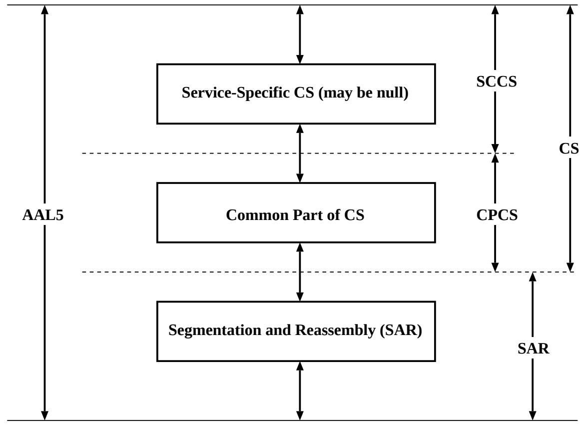

Figure 3.1: Structure of AAL5 layer (from [35]).

3. The adaptive method to recover the clock requires a PLL to determine how the buffer at the decoder end is being emptied. This PLL duplicates the function of the MPEG-2 Systems decoder.
4. Since signaling is being done under AAL5, ATM network interfaces will need to support both types of adaptation layers (AAL1 and AAL5), which makes such a choice expensive.

Transport over AAL5

AAL5 was designed for transporting data traffic with no real-time constraints over ATM. Its convergence sublayer consists of two components: the *Common Part CS* (CPCS) and the *Service-Specific CS* (SSCS). The Common Part CS of AAL5 can make the use of variable length *protocol data units* (PDUs) of size up to 65536 bytes. This means that the size of an AAL5 PDU need not be fixed. The optional Service-Specific CS (SCCS) provides the flexibility of having a special sublayer for different services that need to use AAL5.

The CPCS together with the SAR layer provides all the capabilities to send and receive a Common Part AAL5 *service data unit* (CPAAL5 SDU) from an ATM network. In the case that the SDU is corrupted or lost, an indication is sent to the SSCS (or the service layer if the SSCS is null). However, corrupted or lost SDUs are not recovered as in the case of AAL1 with forward error correction. A corrupted SDU may optionally be forwarded to the SSCS layer, leaving it the responsibility of dealing with the error. A corrupted SDU can be detected by making use of a 32-bit CRC at the end of the SDU and by verifying the length of the SDU from its header field.

In the case of audiovisual services, a Video Audio Service-Specific Convergence Sublayer (VASSCS) was proposed in order to provide the clock recovery and the constant end-to-end delay that MPEG-2 requires [4]. The VASSCS could be designed to support both CBR and VBR MPEG-2 services and more generally other audiovisual services with similar functionalities as the MPEG-2 Systems layer. To implement a constant-delay service, VASSCS should employ a timestamping mechanism and a dejittering buffer at the receiving adaptation layer. Thus, VASSCS duplicates functions that are present at the MPEG-2 Systems layer. At the same time, it is not clear whether adopting this solution justifies the increased cost and complexity by means of better quality. In a recent study with CBR MPEG-2 Transport Streams, Perkins and Skelly [39] showed that with only a small amount of buffering (around 50 Kbytes), the PLL needed at the MPEG-2

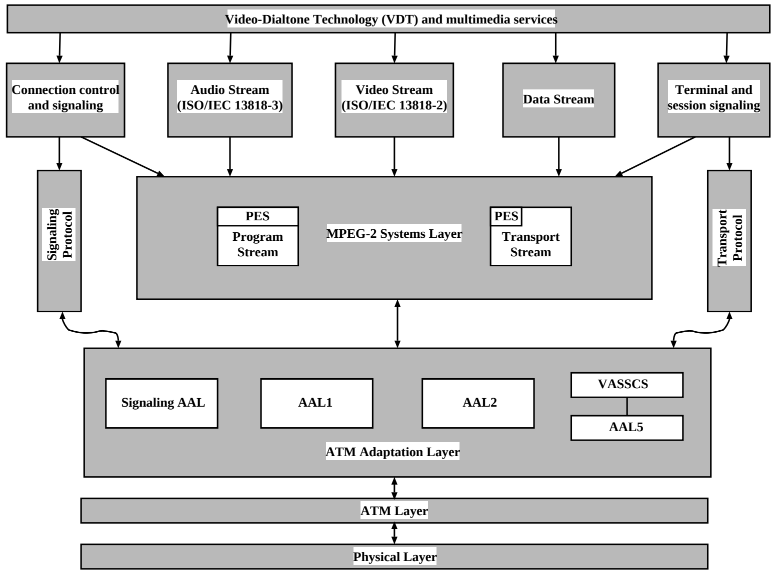

Figure 3.2: Protocol stack for MPEG-2 over ATM (from [9]).

decoder end is still able to recover the source clock when the jitter introduced by the network is limited to 1 ms. This makes the VASSCS solution practically redundant. However, since it may be difficult in practice to negotiate a jitter of 1 ms in the network, the dejittering at the AAL5 layer may still be needed.

AAL5 has several advantages over other alternatives:

1. AAL5 is currently the most commonly used adaptation layer in the industry. AAL5 is being used for encapsulating UNI 4.0 signaling messages and, in most cases, to carry best-effort traffic through the ATM network.
2. Using a null CS, hardware support from the network can be minimized and thus the complexity is moved to the service layer.
3. AAL5 can support VBR MPEG-2 traffic in the future (AAL1 can be used only with CBR traffic).

Despite the above advantages, AAL5 has a number of limitations when used to support real-time applications:

1. The Common Part Sublayer of AAL5 (CPAAL5) does not support forward error correction, as discussed above. However, since congestion losses in an ATM network are likely to occur in bursts, the effectiveness of forward error correction may be limited in any case.
2. Although AAL5 (as it is currently defined) detects errors encountered in a received SDU, it does not forward the damaged SDU to the application layer. This means that, in the case of an MPEG-2 Transport Stream, one or more transport packets may be lost at the service layer, resulting in quality degradation. Besides, having excessive losses makes the error recovery and concealment techniques in MPEG very inefficient. These techniques are typically used to correct synchronization problems at the macroblock or slice level in MPEG-2 video.
3. Different ways of encapsulating MPEG-2 Transport Stream into AAL5 SDUs may affect the jitter of the transport packets, which in turn will affect the system clock recovery process at the MPEG-2 Systems decoder.

Transport Rate (Mbps)	Packing Jitter (ms)
1.0	1.5
2.0	0.75
3.0	0.5
4.0	0.38
5.0	0.4
6.0	0.25
7.0	0.21
8.0	0.19
9.0	0.17
10.0	0.15

Table 3.1: Maximum packing jitter for different transport rates and $N = 2$.

AAL5 is currently the the adaptation layer recommended by ATM Forum specifications for both CBR and VBR MPEG-2 Transport Streams [47, 48]. Although a null SSCS is suggested for the CBR case, ATM Forum is in the process of investigating possible benefits of designing a new SSCS for supporting VBR MPEG-2, to facilitate clock recovery and error control.

3.2 Transport Packet Encapsulation

When AAL5 is chosen as the adaptation layer for carrying MPEG-2 Transport Streams, a scheme must be designed for encapsulating MPEG-2 transport packets into AAL5 SDUs. If each transport packet, which has a fixed size of 188 bytes, is encapsulated in a distinct AAL5 SDU, then five cells will be needed at the ATM layer for each packet (the header and the payload for each ATM cell are 5 and 48 bytes respectively, and the trailer for the AAL5 SDU is 8 bytes). This introduces a significant bandwidth overhead. AAL1 does not suffer from this problem since a transport packet can be carried in a 4-cell AAL1 SDU (only one byte of cell's payload is used by AAL1). To reduce the bandwidth overhead, more than one transport packet need to be encapsulated in a single AAL5 SDU.

Assuming that a single AAL5 SDU will be used to encapsulate N transport packets, there is an inherent packing jitter associated with each of the transport packets carried by the SDU. This introduces a problem when one of these N transport packets contains a Program Clock Reference (PCR) value for system clock recovery at the destination. In that case, the packing jitter will appear as delay variation at the decoder end, and may affect the quality of the clock recovered. To address this problem, two packing schemes for AAL5 were proposed within the ATM Forum: *PCR-unaware* and *PCR-aware*.

In the PCR-unaware approach, the source adaptation layer forms one AAL5 SDU from every N consecutive transport packets without examining the incoming Transport Stream at all. Since all the AAL5 SDUs have fixed size, for the CBR MPEG-2 Transport Stream case, the resulting stream is also CBR and its peak cell rate at the ATM layer can be determined accurately. However, this scheme may introduce significant jitter to the PCR timestamps in the stream, depending on the number of transport packets per AAL5 SDU and the transport rate. Table 3.1 shows the maximum jitter introduced for various transport rates when $N = 2$.

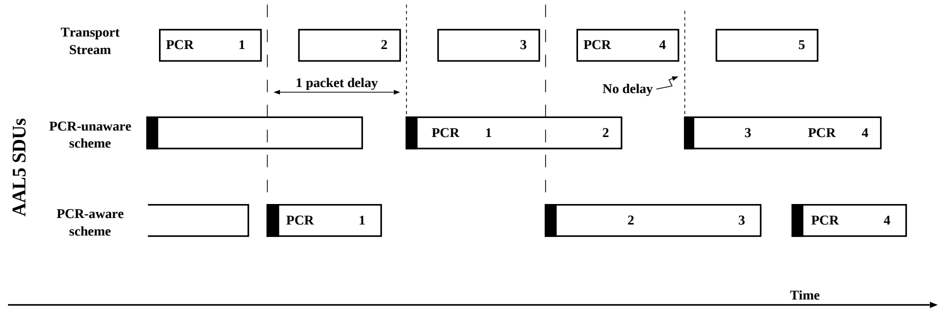

Figure 3.3: PCR packing schemes for $N = 2$.

In the PCR-aware approach, the source adaptation layer forms an AAL5 SDU from N consecutive transport packets except when a transport packet contains a PCR value. On receiving a transport packet containing a PCR value, the AAL5 SDU is forwarded immediately to the SAR sublayer for segmentation. That is, a transport packet containing a PCR value always appears as the last packet in the AAL5 SDU. Thus, the packing jitter for the transport packets that carry PCR values will be essentially zero. However, this approach demands that the source (the encoder or the video server) is capable of detecting transport packets with PCR values. Because the rate of Transport Streams usually exceeds 1 Mbps, this must be done in hardware, adding complexity to the video servers.

The two schemes are contrasted with an example in Figure 3.3. In the PCR-unaware case, the packetization procedure does not examine the incoming transport packets and therefore, the second AAL5 SDU is the result of encapsulating transport packets 1 and 2, whereas the third AAL5 SDU results from the transport packets numbered 3 and 4. The PCR value in the second AAL5 SDU suffers a delay of one transport packet since it has to wait for the second transport packet to arrive before the SDU is formed. However, this is not the case for the third AAL5 SDU since the SDU becomes complete after the transport packet 4 arrives. On the other hand, the PCR-aware scheme completes an SDU if the current transport packet contains a PCR value. Thus, the second SDU is immediately formed as a result of transport packet 1 which contains a PCR value. The third SDU does not contain any PCR values since it contains transport packets 2 and 3. Finally, the fourth SDU is formed and completed by transport packet 4 in its payload without waiting to receive transport packet 5.

Under the PCR-unaware case and for $N = 2$, a transport packet carrying a PCR value may occupy either the first or the second position within the AAL5 SDU. In the PCR-aware approach, when the AAL5 SDU consists of only one transport packet (it must contain a PCR), the SDU has a size of five ATM cells. Since, in the worst case, a PCR may appear in every transport packet, the resulting cell rate may be much higher than that under the PCR-unaware approach. For a CBR MPEG-2 Transport Stream, even if the frequency of PCR values found in the Transport Stream is limited to a reasonable number (50 per second, for example, which is common among encoders that are being used today), the resulting stream is no longer of constant bit-rate. Thus, the peak cell rate of the stream, needed for shaping and call-admission control, must be calculated based on an estimated maximum frequency of PCR values in transport packets. This may result in bandwidth over-allocation and may waste valuable network resources.

The impact of PCR-unaware scheme at the decoder is more memory since the initial locking time for the phase locked-loop at the receiver clock is extended because of the packing jitter, and often poor quality of the recovered clock. This is not the case for the PCR-aware scheme. However, the impact is the additional bandwidth, which may range from 2 to 25% depending on the frequency of PCR samples in the packet stream. Considering the future need for VBR services, the PCR-aware scheme seems the natural choice. The reason is that in the VBR case, the packetization jitter will vary as a function of the instantaneous bit-rate. The largest jitter will be determined by the lowest bit-rate employed in the VBR encoding which may not be known to the decoder a priori. ATM Forum, however, has chosen the PCR-unaware approach ($N = 2$ is the default supported case) as the preferred means of transporting MPEG-2 over ATM [47, 48] for both CBR and VBR MPEG-2 Transport Streams.

When the PCR-unaware scheme is used as the encapsulation scheme, a number of techniques can be used to reduce the effect of the packetization jitter introduced by the scheme [1, 50, 51]. Some key ideas are given below:

1. **Control of the generation of transport packets containing PCRs:** When $N = 2$, PCRs may be placed in even- or odd-numbered transport packets. By controlling the position in the MPEG-2 Transport Stream where the PCR values fall, the packetization jitter may be eliminated, or reduced so that no significant degradation occurs in the quality of the recovered clock at the receiver [1, 51].
2. **Delaying packets at destination:** Packing jitter occurs only when the PCR is in the first transport packet of an AAL5 SDU. This jitter can be compensated at the destination by letting the second packet wait at the AAL until the first packet is transferred to the service layer. This delay can be computed from the current transport rate information [1].
3. **Static jitter compensation in PLL:** Since the amount of packing jitter is known a priori, subtraction of a fixed amount from the PCRs at the destination would eliminate the packing jitter [1].
4. **Restamping at receiver:** By making use of a jitter estimator on the receiver side, the MPEG-2 Systems decoder PLL can be designed to minimize the effect of not only the packetization but also the network jitter [50]. The challenge is to estimate the jitter. The phase difference in the PLL arises from three sources: frequency difference between encoder and decoder, jitter due to network congestion, and packetization jitter at the adaptation layer. The first component is usually small compared to the second and third. Thus, if the magnitude of the resulting error terms at the PLL input crosses a pre-determined threshold, it can be interpreted as being caused by either network jitter or packetization jitter.

3.3 Service Class Selection

Another problem that arises in the transport of MPEG-2 over ATM is in selecting the service class. There are several approaches proposed in the literature for MPEG-2 transport over ATM:

Deterministic Constant Bit-Rate (CBR) approach: In this approach, MPEG-2 is considered CBR in the network and is treated as such. The constant rate to be allocated in the network, called the *effective bandwidth*, has to be either computed in the case of a pre-existing MPEG stream or estimated in a real-time application. Any smoothing necessary to deliver a constant bit-rate stream must be done at the encoder via buffering. If the buffering at the source is limited, we may either use a local feedback scheme to adjust the bit-rate produced by the encoder (Q-factor) or let the buffer overflow. Both approaches result in quality degradation. However, the degradation is different in those two cases. Several methods for the computation of effective bandwidth can be found in the literature [5, 10, 11, 12, 29].

Deterministic CBR without smoothing at the source: In this approach the effective bandwidth is calculated in the same way as in the previous approach, and is used by the network to allocate resources. The difference is that the actual traffic transmitted is of variable rate, so as to gain from statistical multiplexing of different VBR sources. The problem here is that a large amount of buffering may be needed in the switches to exploit statistical multiplexing. If only limited buffering is available, losses will lead to quality degradation. However, recent results [8, 43] show that, in many cases, the effective bandwidth techniques are unable to capture the effect of statistical multiplexing, and hence allocating network resources based on these models would be conservative. Both Choudhury et al. [8] and Shroff et al. [43] verified the limitations of the effective bandwidth techniques. Based on these results, new schemes have been proposed [3, 6, 7, 26] to calculate the bandwidth used for admission control that improve network utilization significantly without the need for large amount of buffering.

Statistical Service with Rate Renegotiation: Two recent schemes have been proposed in the literature based on this approach. Renegotiated CBR (RCBR) [15] is an attempt to combine the simplicity of CBR in terms of bandwidth allocation and call admission control with the advantages of VBR in terms of multiplexing gain. However, the multiplexing gain is not achieved inside the network due to buffer sharing but at the source. The idea behind RCBR is to keep the traffic within the network close to CBR to reduce congestion and jitter. A source under RCBR scheme renegotiates its rate over long time-scales using signaling. Between renegotiation points, the rate is assumed to be CBR and equal to an effective bandwidth value that applies during that time interval. The RED-VBR scheme [52] attempts to eliminate the overload situations in a similar way by using renegotiation points. However, the difference is that RED-VBR builds the renegotiation service on top of a deterministic variable bit rate (D-VBR) service [25, 28], whereas RCBR builds the renegotiation service on top of a constant bit-rate (CBR) service. Thus, RED-VBR achieves better network utilization for the same level of blocking probabilities whereas RCBR offer simpler management and admission control. Both schemes provide a statistical service in the sense that a renegotiation request may be denied. In that case, the source must take proper action to avoid packet losses and the resulting quality degradation. A number of algorithms have been proposed in the literature [14, 20, 34, 41] that perform smoothing on prestored or live video streams and choose optimal renegotiation points to be used by either service scheme (RCBR or RED-VBR).

Feedback-based best-effort service with or without resource reservations: A number of schemes have been proposed for transporting video over a best-effort service where the source adjusts its rate based on available-rate information received from the network periodically (ABR service) [22, 23, 30]. This requires varying the encoding rate at the source adaptively based on feedback information received from the network. Variations of the above schemes have been proposed with and without support for a minimum guaranteed rate. These schemes allow efficient utilization of the network bandwidth, but may result in unacceptable quality for the viewer.

Statistical service without any guarantees: In this case the stream is transported over the network in best effort mode with no feedback controls (UBR service). The quality at the receiver depends on the current congestion level in the network. This is the case with currently available tools for transporting video over the Internet [13, 24, 31].

In selecting a specific type of service for video transport, a compromise must be made between two conflicting requirements: quality-of-service guarantees and network utilization. Deterministic

CBR is certainly the preferred choice for CBR MPEG-2 Transport Streams. Appropriate scheduling disciplines need to be selected for both the ATM end-nodes and the ATM switches over which the CBR MPEG-2 Transport Stream is to be transported. The scheduling disciplines must be able to guarantee not only bandwidth but also low worst-case delays to ensure good quality to the end-user.

For the VBR MPEG-2 Transport Stream case, renegotiation seems to be the most appropriate approach since it attempts to capture the VBR nature of the traffic while maintaining very good quality. An MPEG-2 Transport Stream has a piecewise constant bit-rate [18] that can be used in order to select the renegotiation points. In [17], a new scheme is proposed to propagate changes in transport rate to the decoder end so that the decoder will be able to filter the cell delay variation (CDV) by knowing the actual rate. Although this could be done at PCR boundaries only (at least one every 0.1 sec), in that scheme a *Rate Change Indicator* is proposed to be sent with the Transport Stream exactly when the rate is being changed at the encoder end. A problem with this approach is that it changes the MPEG standard and the adaptation strategy which now have to include this information. However, it is desirable to signal this change of rate information to the ATM network so that it adapts to the new rate of the source. Algorithms need to be defined to compute optimal renegotiation points that do not overload the signaling and maximize network utilization while keeping quality at a high level. The bandwidth requested by the source between any of these points needs to be computed from the MPEG-2 Transport Stream and not from the MPEG-2 video stream since the statistical properties of the video source may change after the processing and multiplexing that occurs in order to obtain the actual MPEG-2 Transport Stream.

In the case of MPEG-2 scalable encoding, combinations of the approaches stated above may work efficiently. For example, the base layer of a video elementary stream may be sent over a deterministic CBR service using its effective bandwidth, whereas the enhancement layer may be sent over an available bit-rate (ABR) service. However, synchronization issues arise in many of these cases and need to be examined very carefully.

Prioritization is another aspect that needs to be considered carefully. Cells carrying MPEG-2 payloads can be assigned priorities based on various criteria. The following are some of the criteria that could be used to determine priorities:

1. Transport packets that contain clock information, i.e., packets containing PCR values.
2. Transport packets from the base layer in the case of scalable MPEG-2 video stream.
3. Transport packets that denote the start or the end of a Packetized Elementary Stream (PES) unit.
4. Combinations of all the above.

The cell priorities can be exploited by the network in controlling not only the distribution of packets that are dropped by the network in case of congestion, but also the bandwidth and delay distribution of the packets. For example, scheduling algorithms could be designed to minimize the delays of certain packet streams to reduce the correlation in the delay distribution. The network, however, must maintain stream continuity even when different priorities are used for cells within the same stream. Typically, the prioritization affects only the buffering decisions in the ATM switches during congested periods.

3.4 Clock Synchronization

Cell Delay Variation (CDV) at the network level is not desirable since it introduces synchronization problems between the source (or the encoder) and the decoder. Several approaches have been proposed for clock recovery and synchronization of MPEG-2 streams in the presence of jitter. The traditional approaches use a PLL to recover the clock from the PCR timestamps transmitted within the stream. The presence of even a modest amount of jitter in this case can adversely affect the quality of the reconstructed clock. Several techniques have been proposed in the literature for improving the quality of the recovered clock. A common technique is to use a dejittering buffer at the receiver that absorbs the jitter introduced by the network. This makes the network transparent to the decoder phase-locked loop. A disadvantage of this approach is that it requires a priori knowledge of the maximum delay variation to avoid overflowing or underflowing the dejittering buffer. In addition, this approach wastes memory by using two separate buffers, the system decoder buffer and the dejittering buffer. Another approach to tolerate jitter at the receiver is to use special pre-filtering techniques to filter the delay variation before the PLL [16].

A third technique to minimize the effects of jitter in the clock recovery process is by counting the time difference between successive timestamps in the packet stream [21]. Although the jitter introduced by the network may be computed on a per packet-basis in this scheme, it requires constant spacing between timestamps in the packet stream, an assumption that may not hold in MPEG-2 Transport Streams. Finally, Akyildiz et al. [1] proposed a simple method to deal with the packetization jitter of CBR MPEG-2 Transport Streams in an ATM network by subtracting a fixed offset from the received timestamps. This scheme, called *Enhanced 2/2 scheme*, deals only with the packetization jitter, and is not designed to correct network-induced jitter.

All the above dejittering approaches attempt to maintain a constant buffer occupancy at the receiver and can therefore be applied to only constant bit-rate streams. In the case of a variable bit-rate stream, constant buffer occupancy is difficult to achieve without knowledge of the rate changes. These rate changes, in principle, can be determined from the PCR values in the stream using their piecewise linearity property [18]. However, changes in the transport rate cannot always be determined exactly from the PCR values. An interesting solution to this problem was proposed by Hodgins and Itakura [17], where a rate change indicator is sent within the stream. The drawback of this scheme, however, is that it requires changes to the MPEG-2 standards. Alternative approaches for clock recovery in variable bit-rate streams include the use of a control system for frequency estimation and adjustment in order to provide constant average delay through the buffer [44]. In the simple case of elementary variable bit-rate streams, another option is to keep a constant number of frames in the buffer instead of constant buffer occupancy. By estimating the frame rate, the problem then becomes a constant rate problem again. However, this cannot be done in the case of MPEG-2 Transport Streams since the arrival rate of the access units of the stream varies through time generally.

A last approach is based on the observation that although the quality of the reconstructed clock is degraded even with moderate amounts of jitter, the jitter does not not cause the MPEG-2 Systems decoder buffer to overflow or underflow [50]. This suggests the possibility of combining the two buffers —the dejittering buffer and the system decoder buffer— and providing a constant amount of dejittering space in the system decoder buffer by subtracting an offset from incoming PCR values. The effects of jitter on clock recovery can be minimized through the use of a jitter estimator to calculate the jitter on a per-packet basis and by *restamping* incoming transport packets containing PCR values based on the estimated jitter. Efficient techniques to estimate the end-to-end jitter found in the literature [33, 32] can be used by the jitter estimator. However, in any case, the network has to be able to guarantee a specific worst-case delay and bound the overall jitter. Bounded low jitter means a less complex PLL at the decoder end which minimizes the cost in a set-top box not only

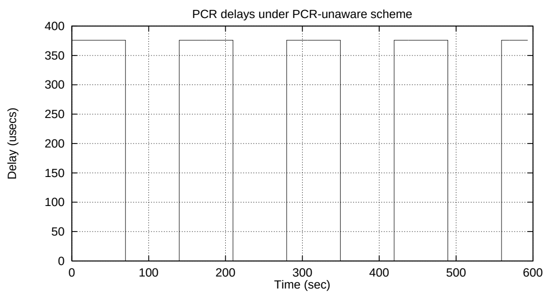

Figure 4.1: Delays experienced by transport packets containing PCR values under PCR-unaware scheme for a 4 Mbps MPEG-2 Transport Stream.

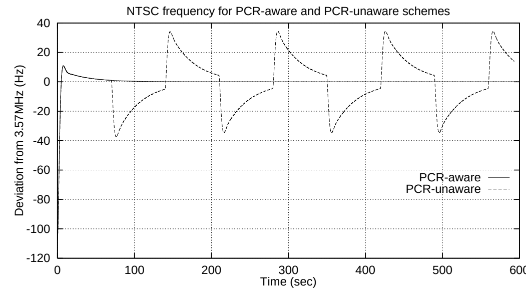

Figure 4.2: NTSC color sub-carrier generation frequency of a 4 Mbps MPEG-2 Transport Stream under PCR-aware and PCR-unaware packing schemes.

because of the complexity of the clock recovery process but also because of less memory required. Thus, the provision of QoS in the ATM network is critical.

Finally, synchronization in distributed playback environments must be imposed using special feedback algorithms as noted in [40]. In those cases, clock information from the corresponding decoders is fed back to the source which then may take appropriate action to ensure that all the decoders are synchronized within a close range.

4 Experiments

To evaluate the impact of some of the issues discussed above, we present some results from simulation experiments with MPEG-2 Transport Streams sent to an MPEG-2 decoder over an ATM network.

We start with a simple experiment to show the impact of the PCR-unaware scheme on the clock recovery at the MPEG-2 Systems decoder PLL. We send a 4 Mbps CBR MPEG-2 Transport Stream through an ATM network under both PCR-aware and PCR-unaware schemes. In this case, the packetization jitter for the PCR-unaware scheme is approximately 376 μseconds. The delay behavior is shown in Figure 4.1. The quality of the recovered clock on the receiver side is heavily degraded in the PCR-unaware case as opposed to the PCR-aware case even though the transport rate is relatively high (Figure 4.2).

In a second experiment, a dejittering buffer is deployed at the destination to minimize the effects of packetization jitter due to the PCR-unaware scheme. We now send a 9.4 Mbps CBR MPEG-2 Transport Stream under PCR-unaware scheme. As shown in Figure 4.4, when dejittering is used, the quality of the recovered clock is perfect with no disturbances compared to the standard case in which the incoming stream is used directly to recover the clock for synchronization. However, the difficulty is that the maximum jitter must be known a priori so that the dejittering buffer can be designed to never underflow or overflow. This may be difficult in some environments.

A third experiment is presented to show that the network itself should provide some level of quality-of-service and that dejittering schemes on the receiver side do not always solve the problem. In this experiment we vary the quality-of-service given by the ATM network by using two different scheduling strategies: the well-known FIFO and a fair-queueing scheduling discipline that provides bandwidth guarantees to the end-to-end sessions referred to as *Frame-based Fair Queueing*

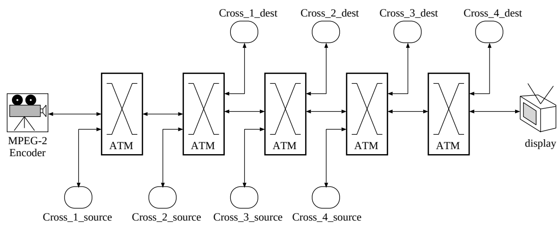

Figure 4.3: ATM network topology used in the simulations.

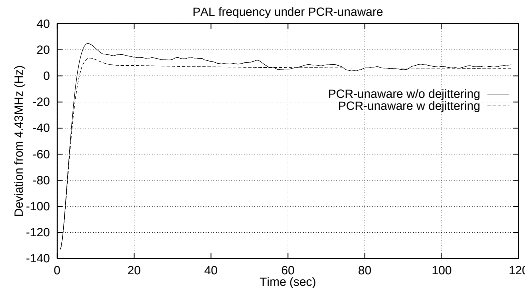

Figure 4.4: PAL color sub-carrier generation frequency of a 9.4 Mbps MPEG-2 Transport Stream sent through an ATM network under PCR-unaware scheme with and without jitter compensation in the MPEG-2 Systems decoder.

(FFQ) [46]. We again send the 9.4 Mbps Transport Stream through the ATM network under heavy load conditions (95%) and obtain the delays and the recovered clock on the receiver side. The ATM network topology used is shown in Figure 4.3. It consists of five cascaded ATM switches. The switch nodes are non-blocking, output-buffered crossbar switches. The MPEG-2 Transport Stream is sent through all the cascaded switches to the display device at the other end. At each hop of the network, the end-to-end video stream shares the network link with cross traffic generated by a set of cell sources. All the cross-connections are between nodes that are connected to adjacent ATM switches.

The maximum jitter experienced by transport packets containing PCR values under FIFO is approximately 22 msecs as shown in Figure 4.5. This results in unacceptable quality of the recovered clock for the FIFO case even with the use of a 10 ms dejittering buffer on the receiver side as illustrated in Figure 4.6. The use of FFQ provided the necessary level of quality-of-service inside the ATM network so that the recovered clock has acceptable quality and minimum disturbances. However, although the quality of the recovered clock was unacceptable under FIFO, the playout buffer dynamics remained almost the same compared to the case when no network load is present (Figures 4.7 and 4.8). This suggests that the provision of clock synchronization and acceptable clock quality is much harder to attain than avoiding underflows and overflows in the playout buffer.

5 Summary and Conclusions

In this article, we reviewed the key issues involved in transporting MPEG-2 streams over ATM networks. The maturity of the ATM standards has made the deployment of such services feasible. Currently, the deployment of MPEG-2 over ATM is realized through the use of CBR MPEG-2 Transport Streams. ATM Forum has chosen AAL5 as the preferred choice for adapting MPEG-2 transport packets into ATM cells, and the PCR-unaware packing scheme for the encapsulation procedure for both the CBR and VBR cases. As shown in the experiments, these choices may introduce problems in the end-to-end quality. It is currently left to the specific implementation to deal with the jitter or losses introduced by the end-to-end path and provide the appropriate levels

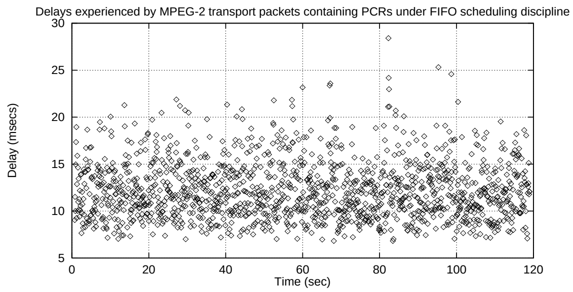

Figure 4.5: Delays experienced by MPEG transport packets containing PCRs with 95% network load under FIFO scheduling discipline.

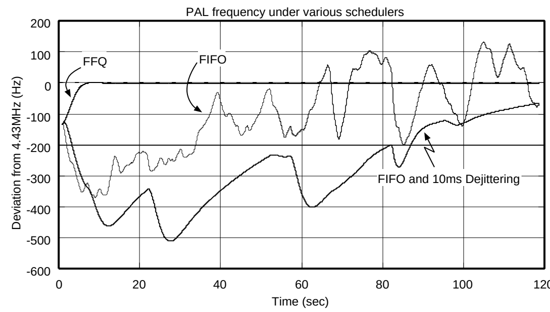

Figure 4.6: PAL color sub-carrier generation frequency with 95% network load under FIFO w and w/o dejittering, and the FFQ scheduling discipline.

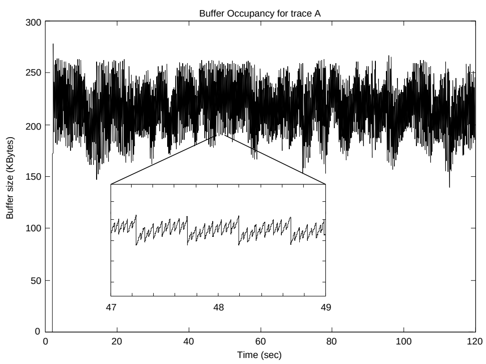

Figure 4.7: Buffer occupancy of the MPEG-2 Systems decoder with no network load.

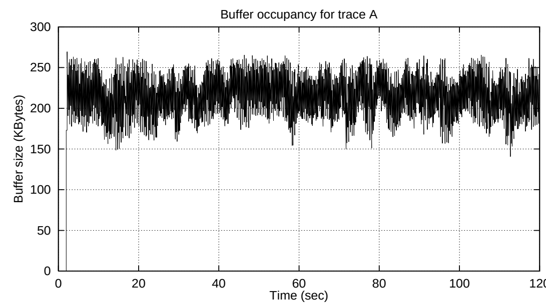

Figure 4.8: Buffer occupancy at the MPEG-2 Systems decoder with 95% network load under FIFO scheduling discipline.

of quality to the end-user. The design of such systems should consider optimizing the end-to-end path as a whole rather than the specific components of the end-to-end path.

As the demand for the support of VBR MPEG-2 over ATM networks grows in order to increase network utilization, new schemes should be identified to maximize both the quality to the end-user and the network utilization. The ATM Forum Service Aspects and Applications (SAA) Sub-Working Group is currently in the process of selecting the type of service required to achieve this. Experiments with new statistical schemes with rate renegotiation for VBR MPEG-2 are necessary to understand how all the issues influence the quality to the end-user under VBR service. The new schemes should be able to capture the bitwise linear properties of the transport rate of MPEG-2 Transport Streams and either statistically estimate or deterministically decide upon the renegotiation points so that network utilization is maximized, while the quality remains at acceptable levels. This is a challenging problem. Also, algorithms that find the optimal or sub-optimal renegotiation points for live sources need to be further investigated.

While many challenges remain in transporting MPEG-2 video streams efficiently over packet networks, the standards infrastructure necessary for supporting high-quality audiovisual services over such networks is emerging fast. Although much of the discussion in this article was centered on ATM networks, many of the issues apply to general packet-based networks equally well. With the anticipated deployment of high-speed national backbone networks capable of supporting video traffic, these issues are likely to assume even greater significance in the future.

References

[1] I. F. Akyildiz, S. Hrastr, H. Uzunalioglu, and W. Yen. Comparison and evaluation of packing schemes for MPEG-2 over ATM using AAL5. In *Proceedings of ICC '96*, volume 3, pages 1411–1415, June 1996.

[2] L. Baptiste. Experiments involving the transmission of layered video over a local ATM network. Master's thesis, University of Toronto, July 1998.

[3] D. D. Botvich and N. G. Duffield. Large deviations, the shape of the loss curve, and economies of scale in large multiplexers. *Queueing Systems, Theory and Applications*, 20(3–4):293–320, 1995.

[4] T. Brunner and M. Schwartz. *Essential Features of a Video and Audio SSCS for AAL5.* ATM Forum, May 1994. ATM94-0404.

[5] Chen-Shang Chang and J. A. Thomas. Effective bandwidth in high-speed digital networks. *IEEE Journal on Selected Areas in Communications*, 13(6):1091–1100, August 1995.

[6] J. Choe and N. B. Shroff. A new method to determine the queue length distribution at an ATM multiplexer. In *Proceedings of IEEE INFOCOM '97*, volume 2, pages 549–556, Kobe, Japan, April 1997.

[7] J. Choe and N. B. Shroff. New bounds and approximations using extreme value theory for the queue length distribution in high-speed networks. In *Proceedings of IEEE INFOCOM '98*, volume 1, pages 364–371, San Francisco, USA, March–April 1998.

[8] G. L. Choudhury, D. M. Lucantoni, and W. Whitt. Squeezing the most out of ATM. *IEEE Transactions on Communications*, 44(2):203–217, February 1996.

[9] S. Dixit and P. Skelly. MPEG-2 over ATM for video dial tone networks: issues and strategies. *IEEE Network*, 9(5):30–40, September–October 1995.

[10] A. I. Elwalid and D. Mitra. Effective bandwidth of bursty, variable rate sources for admission control to B-ISDN. In *Proceedings of ICC '93*, volume 3, pages 1325–1330, May 1993.

[11] A. I. Elwalid and D. Mitra. Effective bandwidth of general markovian traffic sources and admission control of high speed networks. *IEEE/ACM Transactions on Networking*, 1(3):329–343, June 1993.

[12] A. I. Elwalid and D. Mitra. Effective bandwidth of general markovian traffic sources and admission control of high speed networks. In *Proceedings of IEEE INFOCOM '93*, volume 1, pages 256–265, March 1993.

[13] H. Eriksson. MBone: the multicast backbone. *Communications of the ACM*, 37(8):54–60, August 1994.

[14] W. Feng. *Buffering Techniques for Delivery of Compressed Video in Video-on-Demand Systems.* Kluwer Academic Publishers, 1997.

[15] M. Grossglauser, S. Keshav, and D. Tse. RCBR: a simple and efficient service for multiple time-scale traffic. In *Proceedings of ACM SIGCOMM '95*, volume 25, pages 219–230, Cambridge, MA, USA, August 1995.

[16] P. Hodgins and E. Itakura. *The Issues of Transportation of MPEG over ATM.* ATM Forum, July 1994. ATM94-0570.

[17] P. Hodgins and E. Itakura. *VBR MPEG-2 over AAL5.* ATM Forum, December 1994. ATM94-1052.

[18] International Organization for Standardization. *Information Technology — Generic Coding of Moving Pictures and Associated Audio: Systems, Recommendation H.222.0, ISO/IEC 13818-1*, draft international standard edition, November 1994.

[19] International Organization for Standardization. *Information Technology – Generic Coding of Moving Pictures and Associated Audio Information: Video, Recommendation ITU-T H.262, ISO/IEC 13818-2*, draft international standard edition, November 1994.

[20] Z. Jiang and L. Kleinrock. A general optimal smoothing algorithm. In *Proceedings of IEEE INFOCOM '98*, volume 1, 1998.

[21] Y. Kaiser. Synchronization and dejittering of a TV decoder in ATM networks. In *Proceedings of PV '93*, volume 1, 1993.

[22] H. Kanakia, P. Mishra, and A. Reibman. An adaptive congestion control scheme for real-time packet video transport. In *Proceedings of ACM SIGCOMM '93*, September 1993.

[23] H. Kanakia, P. Mishra, and A. Reibman. Packet video transport in ATM networks with single-bit feedback. In *Proceedings of the Sixth International Workshop on Packet Video*, Portland, Oregon, September 1994.

[24] S. Klett. MBONE: videoconferencing over Internet. *CROSSCUTS*, 5(1):10–11, April 1996.

[25] E. W. Knightly. H-BIND: A new approach to providing statistical performance guarantees to VBR traffic. In *Proceedings of IEEE INFOCOM '96*, volume 3, pages 1091–1099, San Francisco, CA, March 1996.

[26] E. W. Knightly. Second moment resource allocation in multi-service networks. In *Proceedings of ACM SIGMETRICS '97*, volume 1, 1997.

[27] E. W. Knightly, D. E. Wrege, J. Liebeherr, and H. Zhang. Fundamental limits and tradeoffs of providing deterministic guarantees to VBR video traffic. In *Proceedings of ACM SIGMETRICS '95*, 1995.

[28] E. W. Knightly and H. Zhang. Traffic characterization and switch utilization using deterministic bounding interval dependent traffic models. In *Proceedings of IEEE INFOCOM '95*, volume 3, pages 1137–1145, Boston, MA, April 1995.

[29] V. G. Kulkarni, L. Gun, and P. F. Chimento. Effective bandwidth vectors for multiclass traffic multiplexed in partitioned buffer. *IEEE Journal on Selected Areas in Communications*, 13(6):1039–1047, August 1995.

[30] T. V. Lakshman, P. P. Mishra, and K. K. Ramakrishnan. Transporting compressed video over ATM networks with explicit rate feedback control. In *Proceedings of IEEE INFOCOM '97*, Kobe, Japan, April 1997.

[31] M. R. Macedonia and D. P. Brutzman. MBone provides audio and video across the Internet. *Computer*, 27(4):30–36, April 1994.

[32] W. Matragi, C. Bisdikian, and K. Sohraby. Jitter calculus in ATM networks: multiple node case. In *Proceedings of IEEE INFOCOM '94*, volume 1, pages 242–251, Toronto, Canada, June 1994.

[33] W. Matragi, C. Bisdikian, and K. Sohraby. Jitter calculus in ATM networks: single node case. In *Proceedings of IEEE INFOCOM '94*, volume 1, pages 232–241, Toronto, Canada, June 1994.

[34] J. M. McManus and K. W. Ross. Video-on-demand over ATM: Constant-rate transmission and transport. *IEEE Journal on Selected Areas in Communications*, 14:1087–1098, August 1996.

[35] D. Minoli. *Video Dialtone Technology.* McGraw-Hill, Inc., 1995.

[36] P. Pancha and M. el Zarki. A look at the MPEG video coding standard for variable bit rate video transmission. In *Proceedings of IEEE INFOCOM '92*, volume 1, pages 85–94, 1992.

[37] P. Pancha and M. el Zarki. Bandwidth requirements of variable bit rate MPEG sources in ATM networks. In *Proceedings of IEEE INFOCOM '93*, volume 3, pages 902–909, 1993.

[38] P. Pancha and M. el Zarki. MPEG coding for variable bit rate video transmission. In *IEEE Communications Magazine*, pages 54–66. May 1994.

[39] M. Perkins and P. Skelly. *A Hardware MPEG Clock Recovery Experiment in the Presence of ATM Jitter.* ATM Forum, May 1994. ATM94-0434.

[40] P. V. Rangan, S. S. Kumar, and S. Rajan. Continuity and synchronization in MPEG. *IEEE Journal on Selected Areas in Communications*, 14(1):52–60, January 1996.

[41] J. D. Salehi, Z. L. Zhang, J. F. Kurose, and D. Towsley. Supporting stored video: Reducing rate variability and end-to-end resource requirements through optimal smoothing. In *Proceedings of ACM SIGMETRICS '96*, volume 1, 1996.

[42] N. Shroff and M. Schwartz. Modelling VBR video over networks end-to-end using deterministic smoothing at the source. *International Journal of Communication Systems*, 7(4):337–348, October–December 1994.

[43] N. Shroff and M. Schwartz. Improved loss calculations at an atm multiplexer. In *Proceedings of IEEE INFOCOM '96*, volume 2, pages 561–568, San Francisco, USA, March 1996.

[44] R. P Singh, Sang-Hoon Lee, and Chong-Kwoon Kim. Jitter and clock recovery for periodic traffic in broadband packet networks. *IEEE Transactions on Communications*, 42(5):2189–2196, May 1994.

[45] P. Skelly, M. Schwartz, and S. Dixit. A histogram-based model for video traffic behavior in an atm multiplexer. *IEEE/ACM Transactions on Networking*, 1(4):446–459, June 1993.

[46] D. Stiliadis and A. Varma. Efficient fair-queueing algorithms for packet-switched networks. *IEEE/ACM Transactions on Networking*, April 1998.

[47] The ATM Forum Technical Committee. *Audiovisual Multimedia Services: Video on Demand Specification 1.1*, March 1997.

[48] The ATM Forum Technical Committee. *Audiovisual Multimedia Services VBR MPEG-2 Baseline Text*, February 1998.

[49] C. Tryfonas. MPEG-2 transport over ATM networks. Master's thesis, University of California at Santa Cruz, September 1996. Available at `http://www.cse.ucsc.edu/research/hsnlab`.

[50] C. Tryfonas and A. Varma. A restamping approach to clock recovery in MPEG-2 systems layer. Technical Report UCSC-CRL-98-4, University of California at Santa Cruz, Dept. of Computer Engineering, May 1998.

[51] C. Tryfonas and A. Varma. Timestamping schemes for MPEG-2 systems layer and their effect on receiver clock recovery. Technical Report UCSC-CRL-98-2, University of California at Santa Cruz, Dept. of Computer Engineering, May 1998.

[52] H. Zhang and E. W. Knightly. RED-VBR: A renegotiation-based approach to support delay-sensitive VBR video. In *Proceedings of the 5th International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV'95)*, Durham, New Hampshire, April 1995.