Interoperability for Professional Video Streaming over IP Networks

By Peter Elmer and Henry Sariowan

Peter Elmer

Henry Sariowan

To encourage the development of interoperable solutions for professional video networks based on the Internet Protocol, the Pro-MPEG Forum has brought together manufacturers, end users, and service providers to develop the necessary codes of practice. This paper provides an introduction to the codes of practice and an explanation of a standards-based approach using Realtime Protocol. It covers protocol details, forward error correction, timing, and jitter considerations. It references evidence of manufacturer implementation within video devices and their use in professional video projects.

Advanced networks based on the internet Protocol (IP) are becoming increasingly commonplace. Support for an ever-increasing range of traffic types sees them playing an important future role in the delivery of professional broadcast content—both realtime and file-based.

The Pro-MPEG Forum has taken a leading role in this area by providing a forum for manufacturers, end-users, and service providers to cooperatively develop interoperable systems for realtime delivery of high-quality program material over wide area networks.

Cost-effective and fit-for-purpose solutions can be engineered by ensuring interoperability between different manufacturers' video devices and the networks to which they connect. This benefit can extend across network domains, ensuring successful delivery of programs between different network providers or geographic regions.

The Forum has made good progress on interoperability for IP networked devices since the publication of "operating points" and Asynchronous Transfer Mode (ATM) device interoperability test results at IBC 2001. Codes of practice for MPEG transport streams and high bit rate studio streams have been developed and supported by manufacturers by adoption in their devices, leading to practical lab tests and public demonstrations.

PRO-MPEG Codes of Practice

Pro-MPEG code of practice (COP) documents for Video on IP have been developed from a practical understanding of issues contributed by members of the wide area networking group. The COPs have been developed from a transmission protocols standpoint, building on existing works published by organizations

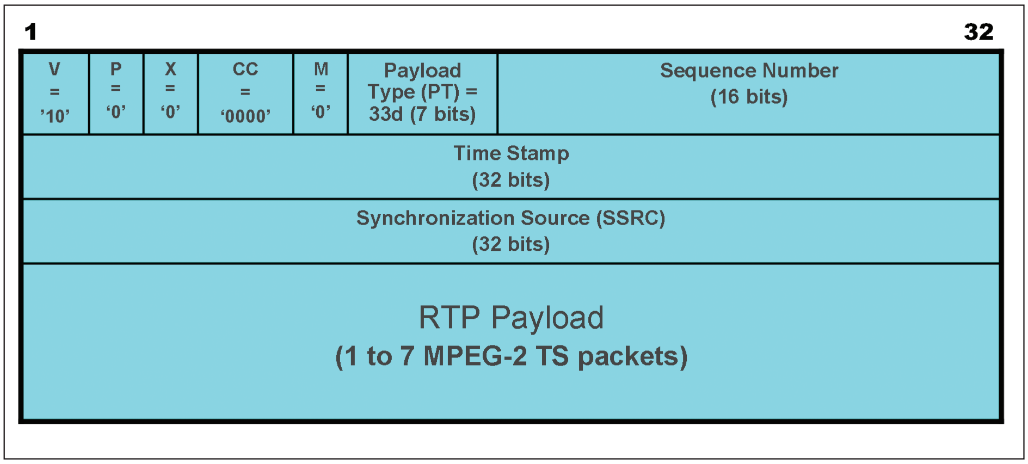

Figure 1. RTP format for encapsulating MPEG-2 Transport Stream within IP packets. V = Version, P = Padding, X = eXtension, CC = Contribution source count, M = Marker, PT= Payload Type.

such as the Internet Engineering Task Force (IETF). The use of the COPs in conjunction with a knowledge of network performance leads to an improved definition of end-to-end service performance.

The published COPs concerned with Video over IP are the following:

COP 3 Transmission of Professional MPEG-2 Transport Streams over IP Networks can be used not only for MPEG-2 video systems, but also for other video formats for which a mapping into an MPEG-2 transport stream is defined, including H.264/MPEG-4 Part 10.

COP 4 Transmission of High Bit Rate Studio Streams over IP Networks covers uncompressed standard-definition video, at 270 Mbits/sec in a way that will not prevent the carriage of compressed formats that use the same framing structure. It is also intended that this document will be applicable to systems running at 360 Mbits/sec and high-definition rates to 1.485 Gbits/sec.

COP 2 Operating points for MPEG-2 Transport Streams on wide area networks is helpful in ensuring the interoperability of MPEG-2 equipment from various manufacturers. The Pro-MPEG Forum has performed several tests of interoperability against them.

These documents collectively promote interoperability by defining encapsulation method and packet format for video payloads and apply constraints to reduce the number of variables. The codes of practice provide guidance to promote interoperability between manufacturers' systems in the following ways:

- Standards-based approach, using Realtime Transport Protocol (RTP) as the baseline.
- Proposed Forward Error Correction (FEC) method, which has acceptance from several manufacturers, will allow protection against packet loss.
- Guidance on latency and packet delay variation issues.
- Clarification on some issues that are not obvious from the RTP specifications.
- Suggested set of operating points to facilitate interoperability testing.

This information should be used by edge device manufacturers to ensure interoperability with other manufacturers' devices.

Protocol Recommendations

Pro-MPEG COP 3 and 4 recommend the use of Realtime Transport Protocol as defined in RFC 3550 as the starting point for encapsulation of video and audio data inside IP packets. The RTP, in conjunction with appropriate RFCs (i.e., RFC 2250 for MPEG-2 Transport Stream and RFC 3497 for HBRSS), specifies how realtime video and audio data should be formatted inside the IP packets, and what additional information should be carried to help the receiver restore the video and audio data from the incoming packets (Figs. 1 and 2). For example, RTP specifies the use of sequence-number and time-stamp fields for the purpose of preserving the order and the playout time of the realtime data. For uncompressed video, the COP 4 extends the header to accommodate for increased sequence number range to extend the roll-over interval and to include additional data, such as scan line number, to improve the ease and robustness of the decapsulation process. In order to minimize the latency caused by packet (de)fragmentation, to simplify the decapsulation processing, and to minimize the impact of packet losses in the networks, the COPs restrict the

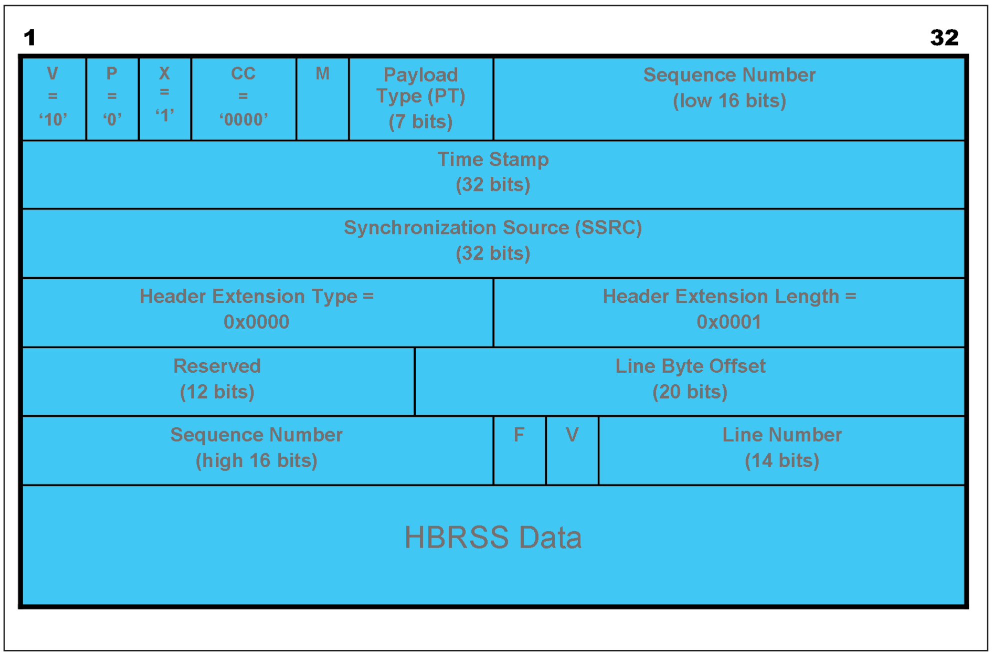

Figure 2. RTP format for encapsulating HBRSS within IP packets. V = Version, P = Padding, X = eXtension, CC = Contribution source count, M = Marker, PT = Payload Type.

size of the IP packet and apply additional constraints. For example, in uncompressed video, all the data inside a packet must be part of only one scan line.

Pro-MPEG's COPs also recommend support for IGMP (Internet Group Management Protocol) as a means for participating in multicast communications within the IP networks. Multicast communications allow a sender to broadcast efficiently a realtime stream to multiple receivers, avoiding the need for stream replication in the network.

The COP also recommends that the Type-of-Service byte in the IP header be configurable by users, so that this byte can be used by the switches and routers to distinguish the priority level of incoming packets and then treats the realtime data at higher priority than other types of data packets. Finally, the Pro-MPEG COP neither requires, nor recommends, any session set-up and tear-down protocols, such as Realtime Streaming Protocol (RTSP) or Session Initiation Protocol (SIP), and leaves them open for manufacturers' implementation. It is also assumed that parameters are fixed for the duration of a session.

Resolving Network Performance Issues

IP networks present a set of issues that have to be resolved in order to make them viable media for transporting broadcast-quality realtime content. Central to these issues is the Quality-of-Service (QoS) that has to be provisioned in the IP networks to ensure that the networks exhibit a predictable performance required by broadcast applications.

Edge devices have to be designed to meet a set of requirements that will complement the network performance requirement. The network performance and QoS issues will be discussed below, as well as how they are addressed by the Pro-MPEG recommendations.

Loss Characteristics and Forward Error Correction

IP networks may cause packets to be dropped for various reasons, the most common being buffer overflow at routers as a result of congestion in the networks.

Packet drops are detrimental to applications and have to be minimized or sometimes eliminated. Thus, the edge devices have to implement a scheme to correct or recover these losses. Because most realtime applications have stringent deadlines on the playout times of packets at the receiver, the FEC technique is preferable to retransmission. FEC involves sending redundant information such that the losses can be recovered from the received data.

Pro-MPEG COP 3 and 4 adopt an FEC method based on IETF RFC 2733, which uses XOR operations performed on a block of packets arranged in a matrix to generate redundant parity packets (Fig. 3). Due to the two-dimensional nature of the FEC matrix, the FEC can recover a burst of losses within an FEC block. The size and shape of the rectangular matrix that forms the FEC block affect the loss recovery capability, the percentage of overhead, and the amount of latency associated with the FEC operation. Thus, it is important to

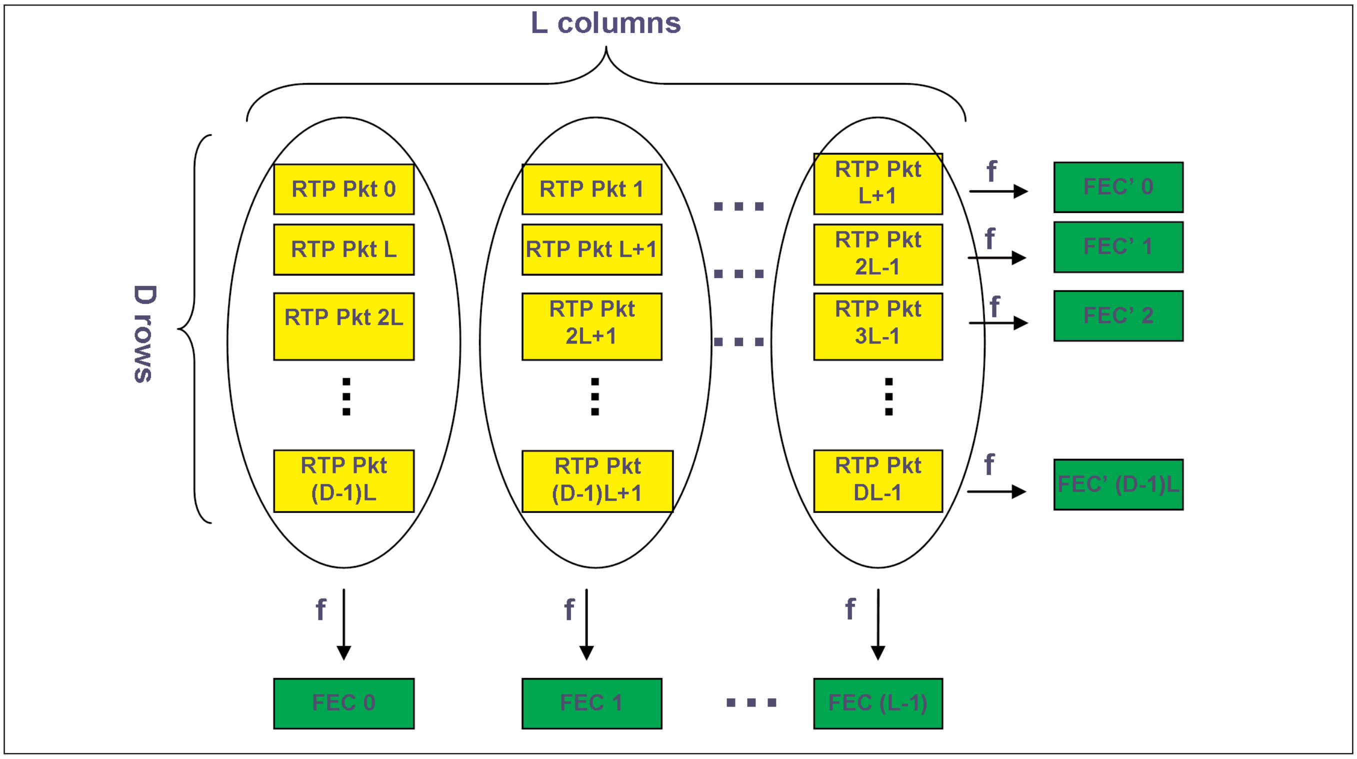

Figure 3. Example of matrix structure for FEC computation.

configure the FEC parameters to match the loss characteristic of the network, as well as to meet the bandwidth and latency requirement of the applications.

Consequently, it is important that packet loss occurrences in the networks are appropriately measured, characterized, and guaranteed by the QoS provisioning mechanisms. Simple network performance metrics such as packet loss ratio may not be sufficient to guarantee that the network loss patterns are appropriately dealt with by the choice of FEC method and parameters. More refined loss metrics, such as Loss Period and Loss Distance as proposed in RFC 3357, will allow the FEC parameters to be configured to match the loss characteristics in the network.

Jitter Characteristic and Buffer Sizing

In addition to packet losses, packets traveling via IP networks may experience a variable delay from the source to the destination. This variable delay is usually called "jitter" and may be caused by several factors, with the most common one being buffering (queue) delay at the switches and routers. Transmission of realtime data such as video, audio, and voice requires that the data be played out at the receiver at regular intervals. In other words, the jitter experienced by the packets has to be "smoothed out" to preserve the timing regularity of the realtime data. The timing restoration process requires the data be buffered at the receiver and then played out at predetermined intervals. The larger the jitter introduced by the networks, the larger the buffer size required by the receivers to smooth out such jitter. The jitter smoothing buffer introduces the undesirable side effect of increasing the end-to-end latency experienced by the applications. Because a realtime application (such as live interviews) may have an upper limit on the end-to-end, or round-trip, transmission delay, the buffer size has to be minimized.

Pro-MPEG COPs recommend a range of buffer size values that has to be supported by interoperable edge devices. However, the manufacturer's equipment should provide flexibility to modify this buffer size as necessary, to handle much better or worse jitter in the networks.

Timing Synchronization

Video and audio signals have very stringent timing requirements that have to be preserved as they are transported over the IP networks. For example, MPEG-2 Transport Streams have to be played out at regular intervals in such a manner that the timing reference in the stream (called the Program Clock Reference, or PCR) does not exhibit jitter of more than 500 nano seconds. This timing preservation requires accurate synchronization between the sender's and the receiver's clocks. Although a jitter smoothing buffer implemented at the receiver will allow video and audio data to be streamed at regular intervals without causing data starvation or overflow, the buffer itself is not a guarantee that the timing integrity of the video and audio streams is preserved according to required standards. The Pro-MPEG COP does not specify a specific algorithm to accurately restore the signal timing from the incoming IP packets; instead, it relies on the edge device manufacturers' know-how and intellectual properties to meet the timing requirement from the standards.

Implementation in Devices and Networks

The Pro-MPEG Forum/Video Services Forum booth on the New Technology Campus at IBC 2004 included a demonstration of practical Video-on-IP interoperability work.

Pre-production units from Path 1 and Thomson successfully interoperated with baseline parameters that excluded the FEC options. Further evolution of the units anticipated over the coming months will provide full support of FEC options in hardware. The demonstration confirmed that manufacturers can commit to, and successfully adopt a common “standard” to achieve multi-vendor interoperability.

Practical evidence of Video-on-IP service implementations has regularly been reported in Pro-MPEG WAN newsletters at previous NAB and IBC events. These include the use of national and international fiber networks to carry video traffic at MPEG-2 transport stream rates of up to 30 Mbits/sec. These have supported conventional unidirectional services and two-way contribution events. Evidence has also been provided of the in-depth investigation and trials undertaken by communication service providers.

Acknowledgments

We would like to take this opportunity to thank the contributors for their significant effort and commitment to achieving great results, with special thanks to Aastra, BT Exact, Path 1 Network Technologies, Tandberg Television, and Thomson for assistance over recent months.

THE AUTHORS

Peter Elmer heads up the broadcast and video technology unit at BT Exact's research, technology, and IT operations division. The unit is responsible for making video technologies accessible to BT's broadcast and broadband service developments. The work of Elmer's unit was previously recognized with the award of the BT Gold Medal for R & D achievement for the delivery of key components of the world's first national digital terrestrial TV network (OnDigital, now FreeView). Elmer's earlier work was on advanced TV systems developing a HDTV system for Europe and a novel system to protect video services against network failures.

Elmer is involved in international broadcasting activities including the Professional MPEG Forum where he chairs the wide area networking group. He is currently involved with the Video Services Forum and is a member of SMPTE. Elmer is also a chartered engineer and a member of the IEE.

Henry Sariowan is vice-president, strategic technology planning, at Path 1 Network Technologies where he provides vision and leadership in Path 1's strategic technology direction. He has also held prior positions as director of systems engineering and systems architect. Prior to Path 1, Sariowan worked at Tiernan Communications designing networking protocols and multiplexing architecture for MPEG-2 high-definition television systems. He has worked at Linkabit Wireless (previously Titan Information Systems) in San Diego, CA, designing a control system for a very small aperture terminal network.

Sariowan received a a Ph.D. from the University of California at San Diego, a Masters of Science from Columbia University in New York, and a Bachelor of Science from the Sepuluh-Nopember Institute of Technology in Indonesia, all in electrical engineering. He is a senior member of the IEEE and an active contributor of the Pro-MPEG Forum and the Video Services Forum.

Presented at the 146th Technical Conference and Exhibition, October 20-23, 2004, Pasadena, CA.