CARRIAGE OF MPEG-4 OVER MPEG-2 BASED SYSTEMS

Ardie Bahraini
Motorola Broadband Communications Sector

Abstract

The MPEG-4 specifications have provided substantial advances in many areas of multimedia technology. In MPEG-1 and MPEG-2 System Specifications referred only to overall architecture, multiplexing, demultiplexing and synchronization of elementary streams. The MPEG-4 specification goes beyond these areas to encompass content description, interactivity, and scene description to name a few. This paper only addresses the overall architecture, multiplexing and synchronization of MPEG-4 content when carried in a system that already supports MPEG-2 Transport Stream.

INTRODUCTION

MPEG-4 is the first digital audiovisual coding standard that expands beyond defining compression algorithms to address emerging computing and telecommunication worlds. MPEG-4 system specifications were intentionally developed to be transport agnostic, enabling MPEG-4 content to be carried over many different transport systems such as MPEG-2, IP, ATM, etc. In particular, MPEG has amended MPEG-2 system standard ISO/IEC 13881-1 to allow carriage of MPEG-4 content over MPEG-2 Transport and Program Streams and this amendment is included in the published 2000 edition of ISO/IEC 13818-1.

The MPEG-2 system standard (ISO/IEC 13881-1) provides two alternatives to carry MPEG-4 content over MPEG-2 Transport Stream (TS). The first scheme is straightforward, and provides the capability for carriage and signaling of individual MPEG-4 audiovisual Elementary Streams (ES) by employing the MPEG-2 system-layer parameters such as PCR, PTS and DTS. This scheme could be used in existing systems that already use MPEG-2 Phy and Transport layers and want to take advantage of the better compression schemes as well as the synthetic video coding tools offered by MPEG-4 part 2 (ISO/IEC 14496-2). MPEG is also extending the MPEG-4 video standard in its specification ISO/IEC 14496-10 (also known as JVT) which will provide significant compression advantage over both MPEG-2 and part 2 of MPEG-4. More details regarding this implementation will be provided in the first part of this paper.

The second alternative defined in the MPEG-2 system layer provides the capability for carriage of MPEG-4 scenes in addition to carriage of MPEG-4 audiovisual elementary Streams. Carriage of this type of content over MPEG-2 Transport Stream follows both MPEG-2 system standards as well as MPEG-4 systems (ISO/IEC 14496 –1) SL_packetize or FlexMux tools specification. This scheme can also be implemented within existing systems in order to provide the MPEG-4 object-based coding and scene composition capability in addition to the better compression that is provided by MPEG-4. The second part of this paper will explore this alternative in more detail.

The last section of this paper will briefly explore how these implementations could be used to enhance current cable systems by migrating from dual carriage of Analog and Digital to an all digital network in order to address future bandwidth requirements as well as providing additional services and features such as HD, VOD and home Gateway based on MPEG-4.

A HIGH LEVEL OVERVIEW OF MPEG-4 DELIVERY LAYERS

MPEG-4 predecessors, namely MPEG-1 and MPEG-2, were designed to address specific systems. For example, MPEG-2 developed Transport Stream (TS) and Program Stream (PS) systems were targeted solutions toward TV Broadcasting and Local Retrieval of content respectively. Hence, the MPEG-2 system was specifically designed to optimize the transport of targeted data and delivery systems by integrating the Sync and Link layers. As a consequence, the MPEG-2 System is not efficient and cannot easily be ported to other mediums and delivery systems without substantial overhead.

MPEG-4, on the other hand, from the beginning was designed to be flexible and independent of under-layer technology such as the delivery system or link layer in order to be adaptable to different delivery systems such as TV broadcasting, IP and ATM systems. To address this separation and be adaptable by different systems, MPEG-4 defined three abstract layers namely: *Compression Layer, Sync Layer and Delivery Layer* as depicted in **Figure-1**. The Compression Layer specifies the encoding and decoding of audio-visual Elementary Streams and is specified by references [2] and [3]. The Sync Layer manages Elementary Streams, their presentation and synchronization information as well as fragmentation and random access information. MPEG-4 Sync Layer syntax is specified by ISO/IEC 14496-1 [4]. The delivery layer specifies the transparent access to other layers independent of delivery technology.

As depicted by **Figure-2**, one could further divide the delivery layer into two sub-layers namely, DMIF (Delivery Multimedia Integration Framework) layer that is specified by MPEG-4 ISO/IEC 14496-6 [*3*], and TransMux Layer that is not specified by MPEG-4 intentionally and is left to the transport technology such as IP, ATM or MPEG-2 to just name a few.

As shown in **Figure-1**, the abstract layer demarcation between Compression Layer and Sync Layer is referred to as ESI (Elementary Stream Interface) and the abstract layer demarcation between Sync Layer and Delivery Layer is referred to as DMIF.

In the following sections, a brief summary of terms related to the MPEG-2 System that are used throughout this paper is offered for those readers who are not familiar with MPEG-2 Sync and Link Layers terminology. Then, the carriage of MPEG-4 elementary streams over MPEG-2 Transport Stream are described as depicted in **Figure-3**. This implementation takes advantage of the MPEG-4 Compression Layer without using other features and functionality provided by MPEG-4. Next, system requirements and architectures are presented for those systems that not only attempt to use MPEG-4 compression, but also intend to use other features of MPEG-4 such as Link Layer and TransMux. This implementation is depicted in **Figure-4**. Finally, in the last section of this paper, the impact of MPEG-4 in the existing Cable TV system is briefly discussed; although a detailed discussion and analysis is beyond the scope of this paper.

A SUMMARY OF THE MPEG-2 TRANSPORT STREAM (TS)

The MPEG-2 system specification ISO/IEC 13818-1, defines two schemes for multiplexing Elementary Streams into a serial bit stream namely, Transport Stream (TS) and Program Stream(PS). The Transport Stream scheme is widely used in transmission of Audiovisual content in CATV today and the Program Stream is used mostly for storage media. In this section only Transport Stream hierarchy is examined since TS is the protocol that is applicable to CATV as noted above.

Each Elementary Stream (ES) contains coded video, coded audio or other data associated with a single program. These streams are separately packetized and formatted into a structure defined by MPEG-2 as Packetized Elementary Stream (PES) as depicted in **Figure 5**. As shown in this figure, each ES could expand into several PES packets. Each PES packet is identified by a stream_id in the packet header. The Stream_id that is associated with each PES packet is defined by MPEG-2 ISO/IEC 13818-1 System specification and identifies the type of steam that is contained in each PES. Each PES is then sliced into Transport Stream Packets that are 188 bytes that include a 4-byte header as shown in Figure 5. The MPEG-2 system specification defines each TS Packet to be identified by a field in the header known as Packet Identifier or PID that is 13 bits. Thus, the payload of each TS packet could contain up to 184 bytes since there are four bytes allocated for the header that include PID in addition to other fields. Each PID is associated with an ES of a service; therefore one program may have one video PID and several Audio PIDs.

Every MPEG-2 Transport Stream multiplex carries a set of tables known as Program Specific Information (PSI) tables. These tables contain information about services, which are present in the multiplex. PSI data includes the following tables: Program Association Table (PAT), Program Map Table (PMT), Conditional Access Table (CAT), and Network Information Table (NIT). Two tables that are relevant to our discussions are PAT and PMT tables. In a compliant MPEG-2 multiplex, there must be only one Program Association Table (PAT) that contains the list of services associated in the multiplex. The PAT in a multiplex associates each service number with a specific PMT PID in the same multiplex. PMT, in turn contains the list of PIDs associated with each service and other associated data. One field of interest to our discussion is the stream_type that is used to associate each component of a service identified by a PID with the type of elementary stream or payload carried within that PID.

CARRIAGE OF MPEG-4 ES VIA THE MPEG-2 TRANSPORT STREAM

This section discusses the encapsulation of MPEG-4 ISO/IEC 14496 audio-visual elementary stream in an MPEG-2 Transport Stream. As noted previously, MPEG has amended the MPEG-2 system standard ISO/IEC 13881-1 [1] to allow carriage of MPEG-4 content over MPEG-2 Transport and Program Streams. This amendment is included in the published 2000 edition of ISO/IEC 13818-1 *Ref*. [1]. According to this amendment, for the carriage of individual MPEG-4 elementary streams, only system tools from MPEG-2 [1] are used. This topology is depicted in**.** As shown in this Figure, MPEG-2 Link Layer and Sync Layer are used instead of MPEG-4 Link and Sync layers. Hence, elementary streams encoded according to MPEG-4 are carried in PES packets as PES_packet_data_types with no specific alignment. In another words, from a system point of view, encoded MPEG-4 audio-visual elementary streams are treated the same as MPEG-2 elementary streams. For example, elementary stream synchronization is accomplished according to MPEG-2 through decoding the PCR in the adaptation layer and the same time base is used to synchronize all the components of a service. This is contrary to MPEG-4 in which each component of a program could be synchronized to a different time base through the OCR.

In addition, Stream_Id values for video and audio elementary stream within the PES header have been defined by this amendment to indicate that PES payload contains MPEG-4 audio-visual elementary streams. As noted in the previous section, Stream_id is encoded in the PES header to indicate to the decoder what compression method is used. The new updated values for Stream_id could be found in table 2-18 of the 2000 edition of ISO/IEC 13818-1 *Ref* [1]. Furthermore, the 2000 edition of ISO/IEC 13818-1 amendment defines stream types that should be encoded in the PMT when MPEG-4 compression is used to encode MPEG-4 audio, and video elementary streams. As noted in the previous section, stream type is used to associate compression used for each component of a service to the PID that is pointed to by PMT table. This information is carried in the second loop of PMT.

New descriptors such as MPEG-4_video_descriptor() and MPEG-4_audio_descriptor() are defined by the 2000 edition of ISO/IEC 13818-1 for defining coding parameters of associated elementary stream. It is worth mentioning that these descriptors do not apply to the MPEG-4 elementary stream if MPEG-4 Link and Sync layers are used. These descriptors are carried in the second loop of PMT and flag to the decoder which level and profile of MPEG-4 compression was used to compress associated Elementary Streams. These descriptors can be found in the 2000 edition of ISO/IEC 13818-1, section 2.6.36 and 2.6.38.

CARRIAGE OF MPEG-4 SCENE VIA THE MPEG-2 TRANSPORT STREAM

This section discusses the encapsulation of MPEG-4 content which may consist of but not limited to: audio-visual, IPMP, OCI streams, Object Descriptor (OD), Scene Descriptor such as BIFS in the MPEG-2 Transport Stream. As specified by the 2000 edition of ISO/IEC 13818-1, these streams are carried in the SL_Packetized stream but use of FlexMux is optional since MPEG-2 offers multiplexing tools. MPEG has amended the MPEG-2 system standard ISO/IEC 13881-1 to allow encapsulation of MPEG-4 both SL_Packetized stream and FlexMux streams in PES packets as well as specifying additional descriptors and other relative fields such as stream_type and stream_id to aid the client side in distinguishing between MPEG-4 content and MPEG-2. Hence, one could use other functionality of MPEG-4 in the existing CATV beyond the improved compression that is provided by MPEG-4. This topology could be represented by Figure 4 and in more detail in Figure 6**.** In the balance of this section the features provided by the latest MPEG amendment for carriage of SL_Packetized and FlexMux in MPEG-2 Transport Streams are presented. First, a brief summary of Sync Layer and FlexMux features and functionality offered by MPEG-4 are provided as they relate to the discussion.

As depicted in Figure 6, the Sync Layer is located between the compression layer and the delivery layer. The Sync Layer provides a flexible set of tools that allow incorporating time base information, fragmentation of access units, and continuity information into data packets. The resulting packetization stream from the Sync Layer is referred to as the SL_packet stream.

The layer below the Sync layer is called the delivery layer that includes the FlexMux. The input to the FlexMux is the SL_packetized stream from the SL as shown in Figure 6. FlexMux is an efficient and simple multiplexing tool defined by MPEG-4 designed for low delay and low bit-rate streams. FlexMux was designed with low overhead since a presentation could have a large number of elementary streams.

Details regarding the Sync layer and FlexMux are beyond the scope of this paper and can be found in ISO/IEC 14496-1 system *Ref*[1].

Figure 6 illustrates the data flow from the compression layer to the Sync layer and then to the Delivery Layer. As depicted in this Figure, the Delivery Layer can be sub-divided into two sub-layers namely, DMIF that is specified by MPEG-4 and the TransMux layer that is not specified by MPEG-4 since different delivery layers already specify this layer. These layers may be any of: RTP, ATM or MPEG-2 just to name few. The following discussion focuses only on MPEG-2 TransMux.

The inputs to MPEG-2 TransMux are either SL_Packetized or FlexMux streams. The responsibility of this layer is to map these streams into the PES and TS structure defined by MPEG-2 with the following constraint: SL_Packetized streams are mapped into single PES Stream such that only one SL_packet constitutes the payload of one PES packet but in the case of TransMux, an integer number of FlexMux packets can be mapped into the payload of PES packets. MPEG-2 defines different a stream_id for PES packets with the payload of SL_packet versus PES packets carrying TransMux streams. The corresponding stream_id should be encoded in the PES packet header.

Furthermore, if SL_packets contain MPEG-4 Object Clock Reference (OCR) then PES packets carrying these SL_packets should have PTS in the PES header and a similar requirement applies to PES packets carrying FlexMux. In case of FlexMux encoded FCR, the timebase in the FlexMux packets should be carried in the corresponding PES header, if present.

Also, certain MPEG-4 streams such as Object descriptor and Scene Descriptor tables can be carried in the MPEG-2 section format. These data are normally static and are used for random access similar to the way in which PSI tables are used in the MPEG-2 system.

Newly defined stream_types are established for the PID streams carrying MPEG-4 content. These stream_types through PMT indicate to the client that the bit stream identified by PID in the PMT is a PES packet containing either SL_packets or multiple FlexMux packets, or that the bit stream contains MPEG-2 section data carrying OD or BIFS commands. Further, a set of descriptors for carriage of MPEG-4 [*Ref*] in MPEG-2

[*Ref*] has been defined. These descriptors provide more information about the stream and are included in the PMT. A list of these descriptors can be found in section 2.6 of MPEG-2 [*Ref*].

WHY MPEG-4

In the early days of cable television all systems focused on delivering clear analog broadcast channels to viewers in fringe areas beyond the reach of broadcast transmitters. As subscriber demand for more selection increased, additional channels were added to the analog line-up. MSOs soon discovered new sources of revenue from reservation PPV, IPPV and premium subscription channels, adding even more channels for these services. It soon became evident that traditional analog CATV system architecture could not support the growing service demand. MSOs were running out of bandwidth and had to upgrade their systems to modern Hybrid-Fiber Coax (HFC) and migrate services to MPEG-2 compressed, digital video transport in order to carry more channels and provide vastly improved picture quality. At the same time ever-increasing competition from the Direct to Home [DTH] satellite networks forced MSOs to seek service differentiation. MSO's began and continue today to offer enhanced features such as HD, VOD, Network PVR, VOIP, Streaming Media and Targeted advertising just to name a few *Ref* [9]. These advanced features pose new bandwidth on both the upstream and downstream HFC plant segments as explained in *Ref* [9]. Thus, once again, bandwidth resources in a typical HFC network are becoming scarce. In order to support these advanced features an MSO can choose to increase the physical capacity of the plant through a costly conversion to an all Fiber network such as FTTH. Or the MSO can optimize the existing plant bandwidth by employing advanced compression techniques such those offered by MPEG-4. Typically, MPEG-4 can provide two to three times better compression than MPEG-2 thereby lowering the effective bit rate without compromising picture quality. Thus a system migration to an MPEG-4 system can accommodate a three-fold bandwidth increase without a costly fiber overbuilds. Given the current investments in MPEG-2, it is not reasonable to retire all this relatively new equipment from service. But through a gradual transition to MPEG-4, as discussed above, it is possible to carry MPEG-4 over an MPEG-2 transport stream and use current or next generation set top boxes that could support this dual mode during a transition period similar to the transition that took place to migrate from analog to digital systems.

CONCLUSION

MPEG-4 is a promising, emerging technology that has been gaining momentum in the CATV industry. This paper has presented different mechanisms to carry MPEG-4 over MPEG-2 Transport Streams in current CATV systems that use MPEG-2 Transport Streams as the transport layer. The first scheme addressed the carriage of MPEG-4 Elementary Streams over MPEG-2 Transport Streams. This scheme enables a CATV system to take advantage of the improved compression offered by MPEG-4. The second part of this paper discussed the carriage of MPEG-4 content over MPEG-2. This scheme enables an MSO to offer the advanced features enabled by MPEG-4 in addition to the improved compression.

Other advanced features such as HD, VOIP, and various Streaming Applications can be supported by migrating to MPEG-4 and freeing the additional bandwidth needed by these applications. It is clear that MPEG-4 is proving to be an industry standard solution for increased efficiency in plant utilization, as well as a much-needed platform for the launch of diverse, evolving interactive services and overall feature enhancements.

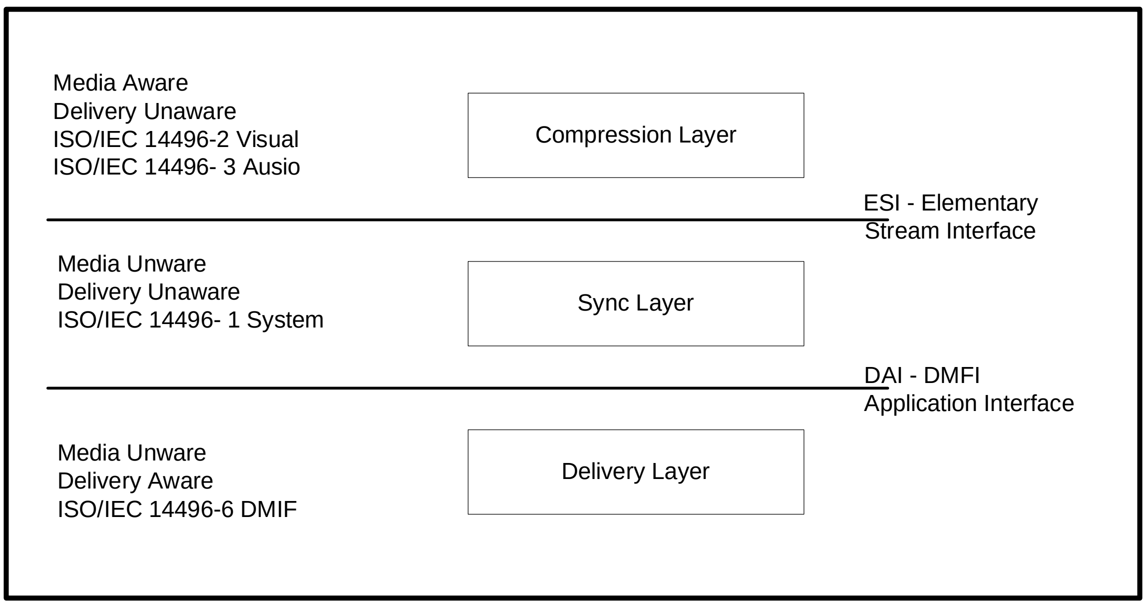

Figure 1 ISO/IEC 14496 SystemArchitecture

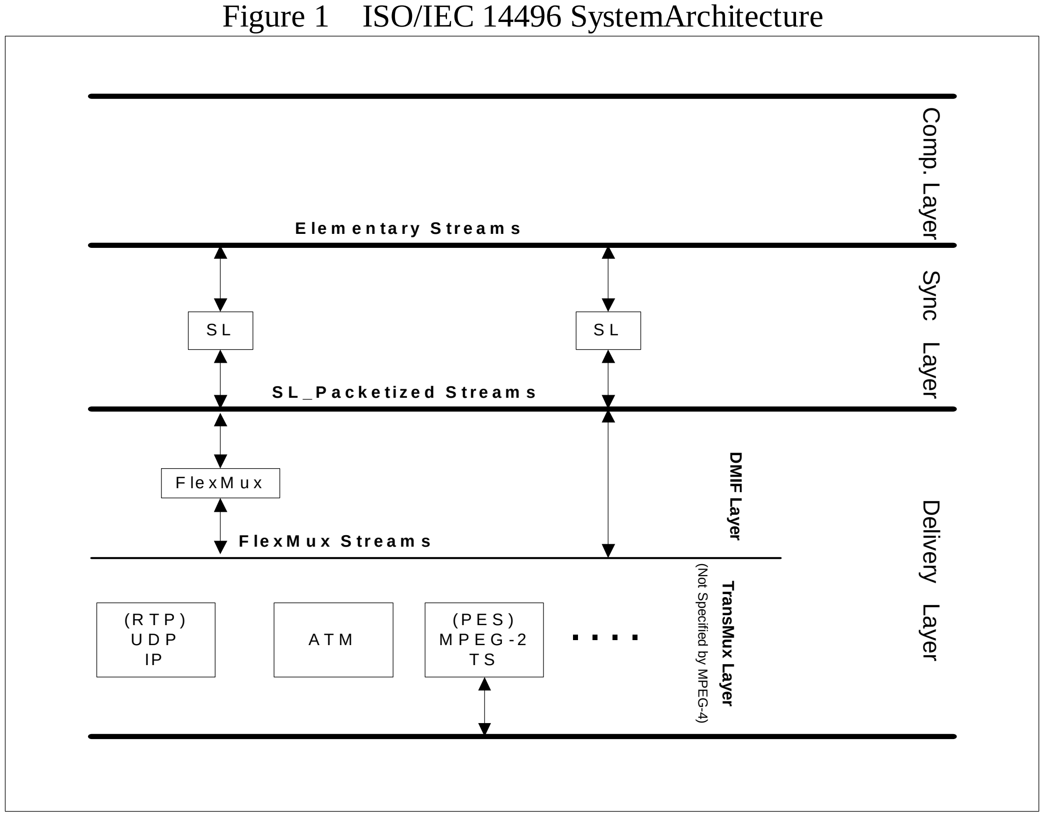

Figure 2 Detailed ISO/IEC 14496 System Architecture

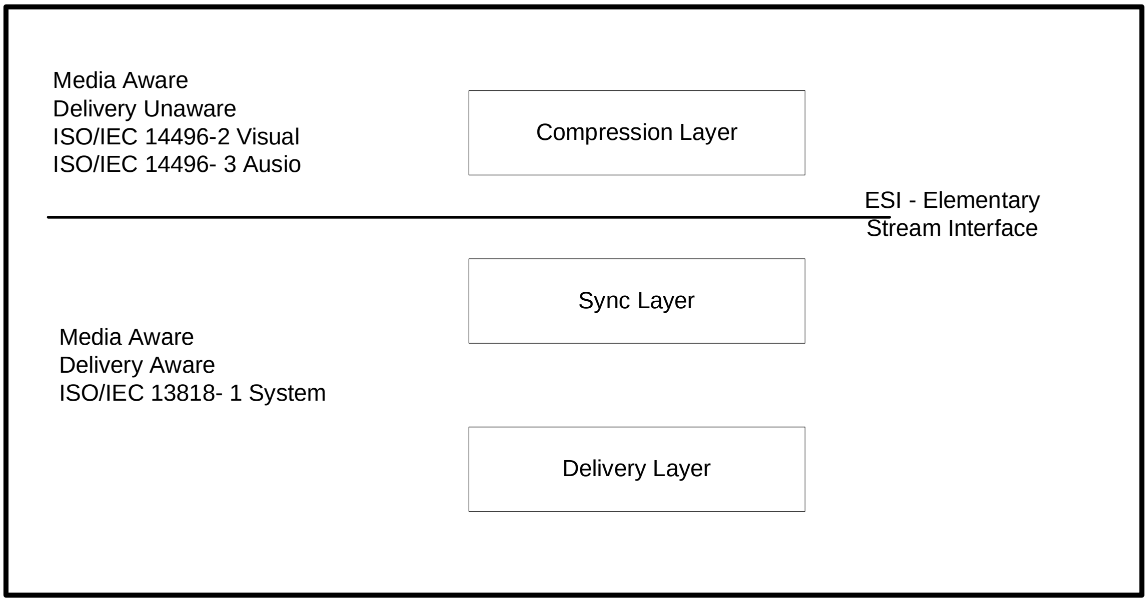

Figure 3 Carriage of MPEG-4 Elementary Stream Via MPEG-2 TS

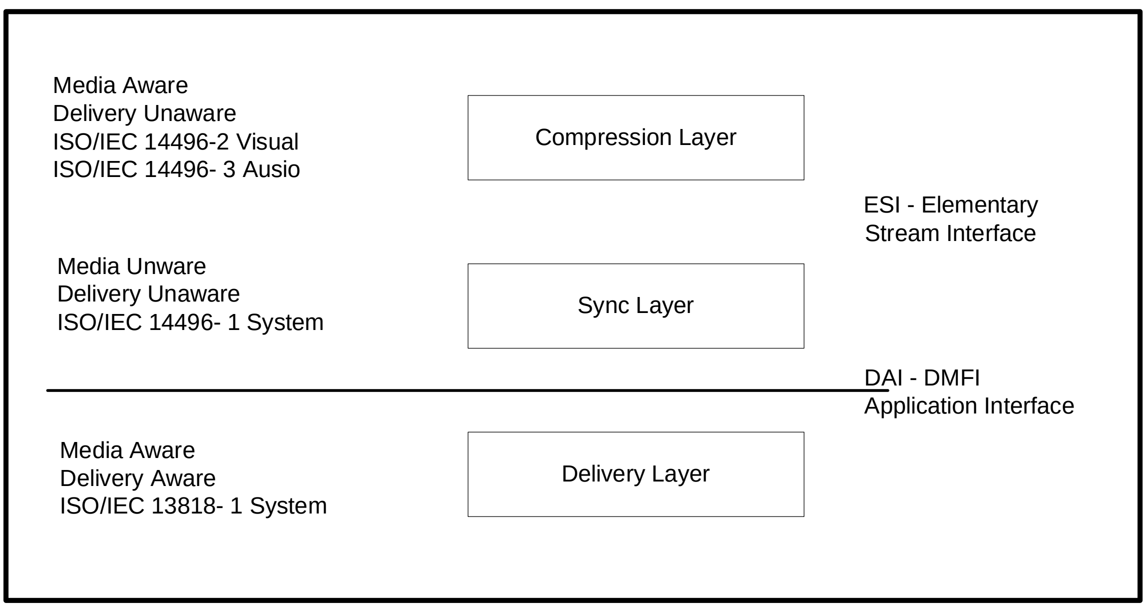

Figure 4 Carriage of MPEG-4 Content Via MPEG-2 TS

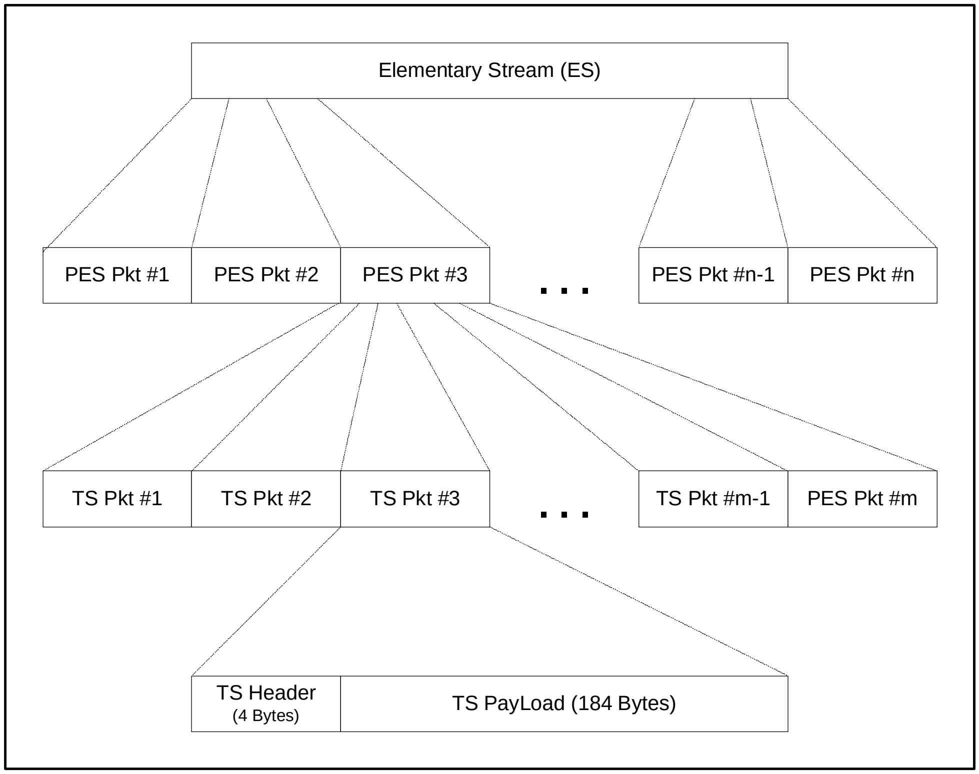

Figure 5 MPEG-2 Link Layer Hierarchy

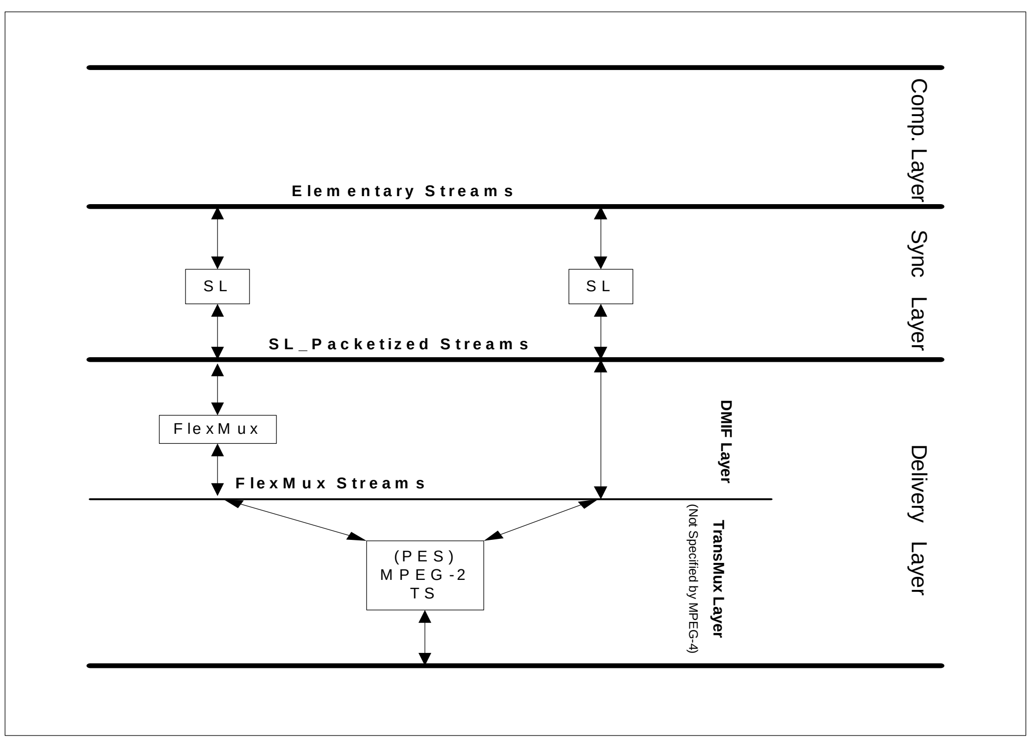

Figure 6 ISO/IEC 14496 System Architecture

REFRENCES

[1] ISO/IEC 13818-1, Information technology – Generic coding of moving pictures and associated audio information: Systems

[2] ISO/IEC 14496-1, Information technology – Coding of audio-visual objects Part 1: System

[3] ISO/IEC 14496-6, Information technology – Coding of audio-visual objects Part 6: Delivery Multimedia Integration Framework (DMIF)

[3] ISO/IEC 14496-1, Information technology – Coding of audio-visual objects Part 2: Visual

[5] ISO/IEC 14496-1, Information technology – Coding of audio-visual objects Part 3: Audio

[6] ISO/IEC JTC1/SC29/WG11 N4668, March 2002, Overview of the MPEG-4 Standard

[7] C. Herpel, A. Eleftheriadis, MPEG-4 Systems: Elemnentary Management

[8] G. Franceschini, The Delivery Layer in MPEG-4

[9] Use of MPEG-4 in Broadband Streaming Applications, Mukta L. Kar, Ph.D. and Nasser Syed, Ph.D. ,Sam Narasimhan, Ph.D. and Ajay Luthra, Ph.D.; NCTA Paper, Year

Author contact:

Ardie Bahraini
Sr. Director, System Engineering
Motorola
Broadband Communications Sector
(678) 957-3233
ardie@motorola.com