ITU Telecommunication Standardisation Sector
Study group 15
Experts Group for Video Coding and Systems in
ATM and other Network Environments

Document AVC 606
Paris 16-18 March 1994
Geneva 14 march 1994

Source: KPN, PTT Research (NL)
Title: A layer model for H.22x, multimedia multiplex and synchronisation for audiovisual communication in ATM environments
Purpose: Discussion

1. Introduction

In the Daejon meeting report AVC-598R, Annex 5 shows a model that has been drawn as a framework for H.22x network adaptation.

In this contribution we present a more detailed layer model in two versions, with or without the MPEG2 Transport Stream layer (TS-layer), as presented infigure 1.

The data exchange between the layers is described in primitives, which are the logical forms and structures of the data. E.g the primitives between packetized elementary stream layer (PES layer) and AAL-CS for PES or TS-layer are PES_packets + timebase information, either parallel or sequential. The primitives and layers are discussed in more detail in the next chapters.

The purposes of this paper are:

- clarify the options and choices for H.22x in combination with MPEG systems
- identify the appropriate interface or Service Access points for a modular implementation, with and without the TS_layer.
- serve as a basis for study on the allocation of functions to layers and the specification of the layers, in particular the AAL-CS and AAL-SAR, with and without the TS_layer.

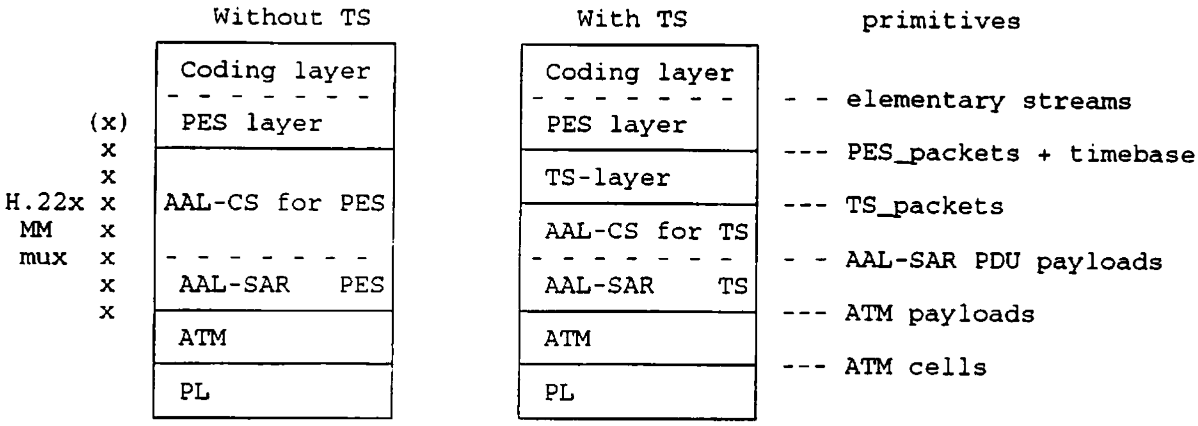

Figure 1. Outline of the layer models

The multimedia multiplexing (MM mux) in the terminal has a generic input from the coding + PES layer and delivers ATM-payloads of 48 octets to the ATM layer. It includes the AAL layer, with CS and SAR sublayers, and may include the MPEG2 TS layer.

The functions of the multimedia multiplex are:

- Mapping of the PES_packets of several streams in ATM cells of one or more virtual channels (one or more VCI/VPI) that have been set up by the terminal. This includes the multiplexing functions segmentation and alignment, and identification of streams by allocation tables.
- Timebase recovery by transfer of timestamps, the timebase is the system time clock (STC)
- Support of buffermanagement in the decoder, closely related to coding layer.
- Synchronous presentation by presentation time stamps (PTS) in the PES_packets. This also is closely related with the coding layer.
- Jitter compensation introduced by the ATM network
- Error protection for bit errors
- Error protection or detection and error handling for cell losses

In H.32x and in H.32y explicitly, H.320 (H.221) streams from existing audiovisual terminals shall be supported as well. The elementary stream multiplex and synchronisation is then according to H.221. The broadband terminal has to support the required circuit emulation through ATM.

I will not further discuss it here.

2. The coding and PES layers: Generic input to the MM muxer

The major benefit of the MPEG system multiplex is the simple timing model with end-to-end synchronisation of the elementary video and audio streams by the inclusion of presentation time stamps referring to a common timebase in the packetized elementary streams. This may be used in H.22x for ATM transmission as well, but H.22x doesn't need to follow the MPEG system multiplex completely.
For audiovisual communication over ATM some of its functions may be better supported in the AAL or ATM layers.

In MPEG also the buffermanagement of the systemdecoder and the decoders of packetized elementary streams is specified by this timing model, referring to the hypothetical system target decoder (STD). The model assumes however a constant transmission delay. In case of timing jitter caused by multiplexing and switching in the network we have to be cautious that the decoder can recover a stable timebase and that the decoder buffers do not overflow.

The timing and buffermanagement model apply to CBR and VBR streams. We will not further discern between CBR and VBR. Invariant support of VBR is mandatory for the multimedia multiplex.

MPEG2 decoders with MPEG2 Transport Stream (TS) interface as the only service access point are expected soon, so interfacing with TS may be a cheaper initial solution for ATM than others. Support by ATM of MPEG2 TS and PS (program stream) or MPEG1 PS is desirable anyway.

2.1 Single program H.262 / MPEG2, with end-to-end synchronization

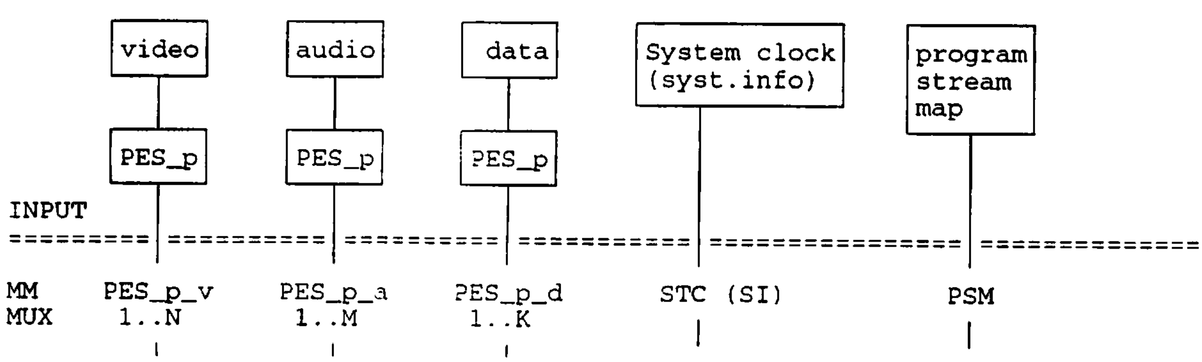

Figure 2. Generic single program input to multimedia multiplexer

PES_p = PES packetizer
PES_p_v/a/d = PES packets with video, audio, data

The generic input of a single program consists of elementary streams of one or more coded video and audio and data streams with the same timebase (STC). These streams are segmented into PES_packets in the PES layer, where the main function of the PES_packet header is the stream identification and support of the end-to-end synchronisation by PTS and DTS. PES_packets have specific stream_id's for each of the N video-streams, M audio-streams and K data_streams in the program. One of the data-streams e.g. could be the MPEG Program Stream directory, if it is available.

In this model the PES_packets with different stream_id's + the STC information are PARALLEL INPUTS to the multiplexer. They may however be in sequential form, with a multiplex on PES_packet level and STC info in pack headers: the Program Stream (PS). In Annex 6 to AVC-598 the output of the PES_layer is proposed as a generic interworking point, but it should be noted that only primitives are specified, unless a sequential PS is created. In general it is not a specification of a bitstream or bitstreams.

Note that during encoding a particuler PES_packet length can be selected and changed according our needs. The PES_packet length may be unspecified unless a multiplexing is performed in the PES layer: in a PS the lengths shall be specified. Unspecified PES_packet length can be advantageous for low delay systems. However, when we retrieve PES_packets from a storage medium the PES_packet length is given. Changing it would require a PES layer recoding function.

The presentation time stamps (PTS's) for audio and video PES_packets and optional DTS's for video PES_packets are referring to the System Time Clock (STC), as a common timebase for the program. The STC information has to be transmitted, to allow the receiver STC" to lock to the source STC. This can be done by transmission of time-stamps in the bitstream.

The time-stamps may be created in the PES_layer as SCR's (=system clock references) if the PES_layer provides a Program Stream. In the general case with parallel inputs the STC is continuously available in some form to the lower layer and the time-stamping is in the lower layer. In the TS-layer the time-stamps are called PCR stamps (program clock reference) and if stamping is in the AAL-CS layer we suggest to call them "STC service time stamps".

The program stream map in figure 2 indicates which streams contain what, e.g. which hierarchical video-layer is in which video stream_id or which language is in which audio stream_id. How this is indicated to the receiver also depends on the type of bitstream that is created.

2.2 Multiple program H.262 / MPEG2 with end-to-end synchronization

If more then one program, with more than one STC, is input to the multimedia multiplex in the form of figure 2, also the relations "which PES_packets belong to what program and to which STC" need to be indicated, in the form of a multi-program map. The information of this multi-program map shall be carried in some form by the TS or its function shall be supported by the AAL.

3. The layers of the multimedia muxer with the TS-layer

In figure 3 a more detailed model of the MM muxer with TS-layer is presented.The TS-layer supports following functions:

- Mapping of the parallel input packetized elementary streams in fixed length transport packets (TS_packets) with size of 188 byte, by the TS_packetizers (TS_p in figure 2).
  This length was chosen as a compromise to meet both the blocksizes of favourite encryption algorithms and the ATM adaptation (188 = 4 x 47). Each transport packet has a specific packet identifier, the PID. Each stream_ID has a unique PID and the start of PES_packets is aligned with TS_packet payload. A stuffing mechanism is included to stuff in the TS_packet where a PES_packet ends.

- Time-stamping of the STC: TS_packets with a specific PID, the PCR_PID, carry program clock reference timestamps in their adaptation_field. There is one PCR_PID per program.

- Transfer of program specific information, like program stream map, in PSI_tables, which are transported in one or more PID's. It provides the program map information to the receiver, e.g. which is the PCR_PID, which PID's belong to the program and what is in them. Reconstruction of the program map table starts from the program association table in PID_0.

- If encryption is applied, conditional access information (CA tables) is carried in PID_1.

- Detection of lost TS_packets in a TS_stream or duplicated transmission of the same TS_packet for error protection is supported by a TS_packet sequence counter, the continuity counter.

- Two transport priority levels for TS_packets can be set by the transport_priority bit.

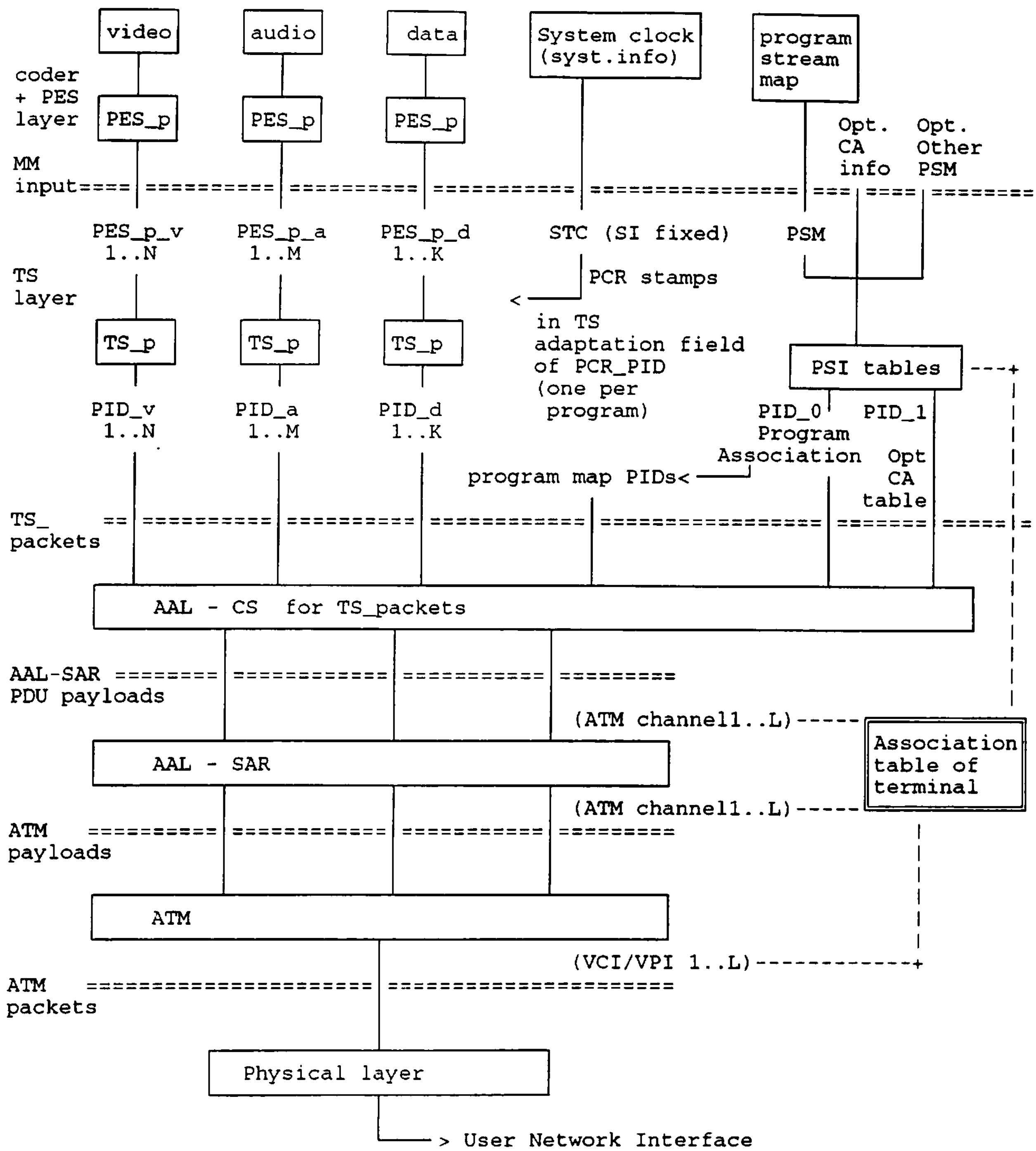

Figure 2. Logical diagram of multimedia terminal with TS layer included

In the general case of our model, the TS_packets are primitives. TS_packets with different PID may be parallel or sequential. An MPEG2 transport stream is a sequential bitstream of TS_Packets.

MPEG2-TS is an agreed interface with major initial support for MPEG2 decoders, especially in broadcasting applications. It is a general interworking point.

The AAL-CS for TS_packets may support the following functions:

- Jitter compensation introduced by the ATM network if the jitter is beyond the margins that can be accepted by the TS and higher layers. This is implementation and application dependent. E.g. for regeneration of a colour subcarrier a very accurate STC recovery may be needed. MPEG doesn't specify acceptable jitter margins, the STD model specifies a zero transmission time.
  Jitter compensation of at least the PCR stamps may be needed in the AAL-CS if the implementation of the TS and higher layers isn't capable to handle jittered PCR's.
- Error protection for bit errors
- Error protection or detection and error handling for cell losses

In our model, one or more ATM virtual channels may be used for the multimedia transmission. The major advantage of using more than one VC is that the network can discern the elementary streams or sets of elementary streams that are in separate VC's in the ATM layer. The selection of programs and elementary streams can be made in the ATM layer and the user will receive only his selected VC's.

The consequence of applying more than one VC is that a form of program and elementary stream association has to be supported by the ATM terminals. In case of TS_packets this can be a simple PID to VCI mapping table. Further study and discussion is needed to consider the impact on call-setup and the control plane.

3.1 Discussion of AAL-SAR for TS_packets

If no interleaving of cells for cell error protection is applied in the AAL-CS, the AAL-SAR sublayer will typically map a TS_packet aligned into 4 subsequent ATM payloads (192 octets), leaving 4 octets per TS_packet for support of the AAL-CS functions.

In case of AAL type 1 the AAL-SAR PDU is 1+47 octets. The 1 octet AAL-SAR PDU header contains cell sequence numbering for detection of cell loss or misinsertion, as well as the CSi bit which may be used for TS_Packet start indication and for the SRTS clock recovery method if a common network clock is available to sender and receiver. Note that the sequence numbering is independent and different from the TS layer. The TS_packet priority bit can be used to set the priority bit in the ATM cells.

AAL type 1 doesn't include TS_packet oriented forward error correction (FEC). Bit error protection per TS_packet isn't possible in this structure, as the AAL1 fully occupies the 1 octet AAL-SAR PDU header and no CRC bits for the payload can be added in the 4 cell structure. Giving up the TS_packet alignment or the 4 cell structure will reduce the efficiency of the multiplexing.

AAL5 may be more suitable, but it would be better to have an AAL that is dedicated to the MPEG TS_packets.
This AALx for MPEG TS_packets would have an AAL-SAR PDU payload of 188 octets and AAL-SAR PDU header of 4 octets. These 4 octets should at least support FEC for bit errors and the SRTS timing recovery method if needed in the AAL.

A new proposal for such an AAL, was recently submitted as AAL6 to the ATM_forum and has been published in an E-mail bulletin [Doc. ATM_Forum/94-0128]. In this proposal the 4 octets per TS_packet are used to support:

- Timebase correction
- Error correction for bit errors in the TS_packets

No cell sequence numbering is included, cell loss implies partial loss of the TS_Packet and the whole TS_packet is considered as lost in the TS layer.
The proposed time base correction seems highly complicated for general ATM exchanges. Forthcoming ATM switches certainly do not support this function and I doubt if it is feasable for high-speed switches anyway:

> It demands a local time stamp compensation for the variable queuing delay in every switching element. This function would better be allocated in the ATM-layer anyway.

Since the current ATM layer doesn't provide the desired jitter measurement data, clock recovery and jitter compensation shall be performed self-supporting by the higher layer, either AAL or TS-layer.
An SRTS method as in AAL1 may be a solution, if a common network clock is available to sender and receiver. An adaptive clock recovery method based on buffer fulness would be suitable for CBR only and therefore not suitable for general MPEG2 signals.

The subject of jitter certainly needs more discussion with regard to MPEG systems: if an MPEG stream is remultiplexed, the PCR time-stamps are supposed to be corrected, so that they represent the correct STC in the output bitstream. This will certainly simplify the clock recovery in the receiver, but the bitstream is still jittered. Additional buffer capacity in the receiver is needed beyond the System Target Decoder values to handle this jitter.

Question: why not specify jitter margins in the standard that every decoder
shall meet to conform to the standard?

With a PCR frequency of at least 10 Hz and a sufficiently accurate free running receiver clock a PLL with narrowband loopfilter may serve for clock recovery and jitter indication in the receiver. An example is found in informative Annex-B of the MPEG2 Systems Committee Draft (ISO/IEC 1-13818 CD).
Evidence is needed if an additional clock system in the AAL beyond the one in the MPEG systems layer would really be needed.

4. Multimedia multiplex without the TS-layer

Use of the TS-layer in combination with ATM transmission is redundant, since functions like cell or packet delineation are already available in ATM.
Functions like sequence numbering and priority indication are better performed with the ATM packets being the real "transport packets", than with the 188 octet TS_packets. Therefore an alternative solution without TS_layer may be worthwile.

In figure 3 the multimedia multiplexer accepts the PES_packets and STC information as primitives. The functions of the TS layer are now integrated in the AAL. Therefore also the program specific information (tables) need to be multiplexed in this H.22x specific AAL. The clock information may be stamped directly in the AAL-SAR PDU headers, if needed by an SRTS like method.

This version without TS_layer avoids the overlap in transport layer functions and offers higher flexibility in the design of AAL, with the possibility of better protection and handling of cell errors.

The burden of this approach is the current lack of a widely accepted common interface point like the MPEG transport stream. A bitstream specification of the MM input as interworking point is urgently needed, e.g. as a "parallel program stream".

As an inbetween solution it has been suggested that TS_layer may still be used, but TS_packet header be stripped for the functions that are not needed in ATM. E.g. the sync_byte has no function if structured data transfer of TS_packets is supported by the AAL.

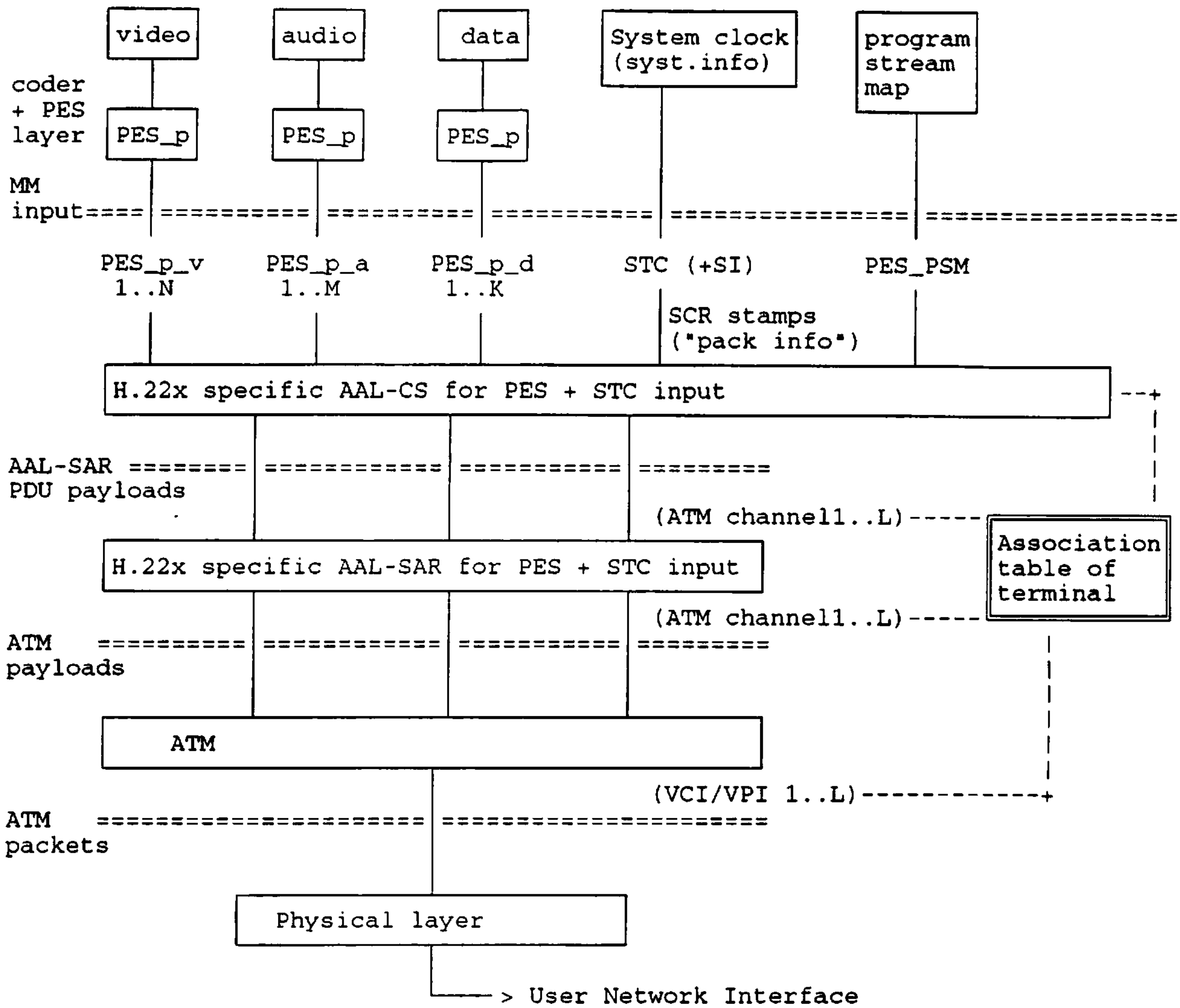

Figure 3. Logical diagram of multimedia terminal without TS layer

5. Conclusion

Some options and possible function allocations for H.22x have been introduced to serve as a basis for discussion.