

Telecommunication Standardization Sector
Study Group 15
Experts Group for ATM Video Coding
(Rapporteur's Group on Part of Q.2/15)

Document AVC-463
March 23, 1993

SOURCE : Japan
TITLE : Consideration of MPEG-1 Systems as multimedia multiplexing for audiovisual communications
PURPOSE : Discussion
Relevant sub-group: Syetem

1. Introduction

It is our objective that one integrated audiovisual terminal can be used for not only conversational service but also retrieval and broadcasting services via ATM networks. For simplicity of integrated terminal, it is preferable that different services use a common multimedia multiplexing method if possible [1] *1. This document discusses what would happen if MPEG-1 Systems are used as a multimedia multiplexing method for conversational services, assuming that MPEG-1 Systems are also used as MPEG-2 program mux.

*1; There is another opinion that it is not a problem even if different services use their own multiplexing methods because remultiplexing is not so difficult.

2. Case study for high quality videoconferencing service

A case study has been carried out for 5 Mbps high quality videoconferencing service in which MPEG-1 Systems are used for multiplexing video, audio, data and control information (Annex 1). When video and audio allow 9 msec delay for multiplexing, overhead of packet header is 0.5% for 4.8Mbps video and is 18% for 64kbps audio. Total overhead of 0.9%, however, is sufficiently small as audio bitrate is low.

Remaining problems are listed in Section 4 Discussion points.

3. Impact of cell losses

The impact caused by cell losses in multiplexed signal is discussed in Annex 2. When MPEG-1 Systems are used for multimedia multiplexing, video decoder should be prepared for not only transmission errors due to cell losses but also packet loss which is equivalent to consecutive cell losses. When the packet interval of a particular medium is n cells, packet loss probability of that medium is $1/n$ of cell loss provability. When a packet loss occurs, error indication to video and audio decoder is impossible. *2

*2; Sequence number for each stream_id packet is necessary for such indication.

4. Discussion points

When we consider MPEG-1 Systems as a multimedia multiplexing method for high quality videoconferencing, the following points should be discussed.

(1) Constraint on number of packets per second (Annex 1)

The constraint on number of packets per second (currently 300 in CSPS) should be increased if multiple audios are considered. This is required to reduce audio packetizing delay. On the other hand video coding causes longer delay than this audio packetization delay. Therefore if the only purpose of user multiplexing is for lip synchronization, the number of packets per second should be evaluated comparing video coding/decoding delay with the sum of audio coding/decoding delay and packetizing delay.

Discussion Point;

Which is necessary; lip synchronization or low delay for audio?

(2) Correspondence between medium and stream id (Annex 1)

Each medium (such as audio, video, data) in multimedia applications may have different number of channels and different coding laws. The number of elementary streams for each medium and coding law for each elementary stream should be known to the decoder a priori or by negotiation in the call set-up process.

Discussion Point;

Are stream_ids standardized dependently on or independent of applications? If stream_ids are application independent, video service integration becomes easier but we need many bits for stream_id. Another possibility will be that the correspondence between stream_id and coding law is conveyed in the terminal-to-terminal control channel

(3) Implementation

MPEG-1 Systems have flexibility for packet length, number of elementary stream etc..

Discussion Point;

Does it cause any implementation problem in case of high bitrate real time communication?

(4) Impact of cell losses (Annex 2)

Packet loss without error indication occurs. The probability of packet loss is $1/n$ of cell loss probability when the packet interval is n cells.

Discussion Point;

Do packet losses without error indication cause a new problem which has not been addressed?

(5) Rate control

Total bitrate control for multiplexed signal comprising several media has not been examined under VBR UPC.

Discussion Point;

When UPC is applied to a multiplexed bitstream, is video coding controllable? The problem may become complicated if multiple VBR sources are involved (e.g. VBR audio and VBR video, two VBR videos).

5. Conclusion

A case study has been carried out when MPEG-1 Systems is used as a possible multimedia multiplexing for high quality videoconferencing service. Several discussion points are clarified.

It should be urgently discussed and decided what changes we should require to MPEG-2 program mux.

[1]Japan; Multimedia multiplexing method for audiovisual communication, AVC-270, October 1992

A Case Study of Multimedia Multiplex for High Quality Videoconferencing

1. Purpose

As a typical example of real time audiovisual communications, we study videoconferencing to find any problems when the ISO/IEC 11172-1 method is applied for multimedia multiplexing.

Specific parameters are chosen, a syntax set is described and performance is evaluated in terms of delay and transmission efficiency.

2. Design conditions

2.1 Number of elementary bitstreams and bit rate

Medium	Bit rate (kbit/s)	Coding law	Note
Audio (A)	64	G.722	
Video (V)	4800	H.26X/MPEG-2	Synchronous with audio
Data (D)	64	JPEG high resolution still pictures	- Asynchronous with audio - Data channel is assumed to be always open
End-to-end control (C)	16	H.24X	- Asynchronous with audio - For in-channel negotiation, control, indication
Encryption control (E)	-	H.???	Not considered in this study for simplification
End-to-network control (S)	-	Q.???	- Dedicated VC is used - Not considered in this study

2.2 Network

- CBR at 5 Mbit/s
AAL: Type 1, hence the signal at the interface between multimedia multiplexer / demultiplexer and AAL is a bitstream with constant bit interval of 1/5 μ s.

2.3 Multimedia multiplexing

Constrained System Parameter Stream of MPEG-1 is applied; number of packers per second is not more than 300. Each packet contains a single medium.

2.4 Packetization delay

We consider packet multiplex of audio and video only to analyze packetization delay. Since audio bit rate R_a is much lower than video bitrate R_v , audio determines the packetization delay (see FigureA.1).

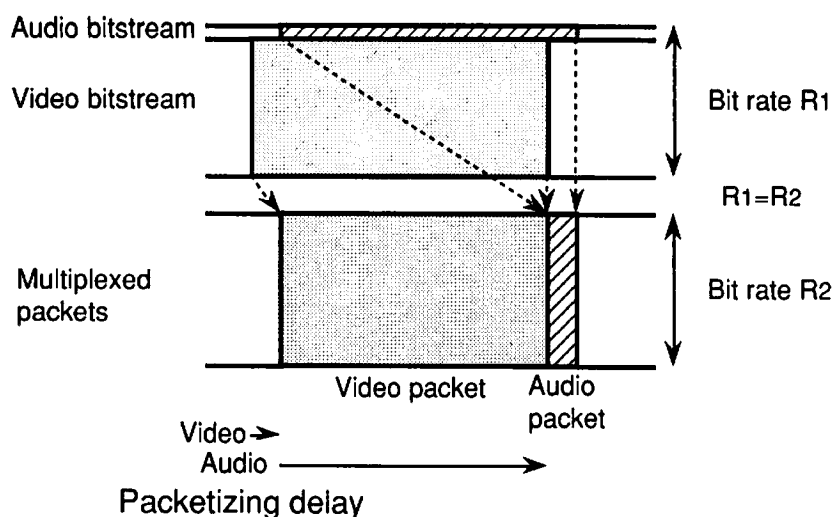


Figure A.1 Packetization delay

The following is observed from this illustration;

- Audio packetization delay is approximately equal to the packet interval.
- Since video bit rate is much higher than audio bit rate, number of video packets can not be less than that of audio.

3. Design results

3.1 Packet assignments

Total 300 packets per second are assigned to each medium. Assuming 100 ms packetization delay is acceptable for Data or End-to-end Control, 10 packets per second is assigned to each of them. The remaining 280 packets are equally divided for Audio and Video.

Medium	Bit rate (kbit/s)	Number of packets per second	Average packet length (byte)
Audio	64	137.93	58
Video	4813	137.93	4361.8
Data	64	10	800
Control	16	10	200
Header (pack, packet)	0.4 for pack 42.6 for packet	1 pack/0.7 sec 296 packets/s	36/pack 18/packet
Total	5000	295.8	-

3.2 Packetization algorithm

- As soon as 58 bytes of the audio signal are stored in the buffer, an Audio packet is sent.
- When 200 bytes of the end-to-end control signal are stored in the buffer, a Control packet is sent just after the earliest available Audio packet.
- When 800 bytes of the data signal is stored in the buffer, a Data packet is sent just after the earliest available Audio Packet. If a Control packet is also available, the Data packet is sent just after the Control packet.
- Video signal is sent by forming a packet occupying between the end of Control or Data packet and the beginning of Audio packet.
- If the buffer approaches underflow, a padding stream packet of appropriate length is sent.

Multiplexed packets are shown in Figure A.2.

```

0000000000111111111122222222223333333333444444444455555555556666666666777777
0123456789012345678901234567890123456789012345678901234567890123456789012345
! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! ! !
PSADVAVAVACVAVAVAVAVAVAVAVAVAVAVAVADVAVAVACVAVAVAVAVAVAVAVAVAVAVAVADVAVAVACVAVAV

```

- P: Pack header (once every 0.7 sec)
- S: System header (once every 0.7 sec)
- A: Audio packet (137.9 packets per sec)
- V: Video packet (137.9 packets per sec)
- D: Data packet (10 packets per sec)
- C: Terminal-to-terminal control packet (10 packets per sec)

Figure A.2 Packet order

3.3 Syntax

Codewords are show for pack header, system header, Audio/Video/Data/Control packets in Appendix.

4. Evaluation

4.1 Transmission efficiency

Medium	Header length(byte)	Information (byte)	Transmission efficiency (%)
Audio	13	58	81.7
Video	22	4361	99.5
Data	13	800	98.4
Control	13	200	93.9
Pack overhead	36	437464	99.99
Overall			99.14

Overhead for pack and packets is 0.86% in total, hence the overall transmission efficiency is sufficient.

4.2 Delay

Delay due to the packet multiplex is 9.9 ms, while packet demultiplexing is negligible. This 9.9 ms delay is considered to be an allowable limit compared with delay of other elements such as audio and video coding. If multiple audios are employed, e.g. one for public talk and another for private talk, the packetization delay increases proportionately to the number of audio channels. Hence the limit of 300 packets per second should be increased at least to twice this value for real time communication.

Note - Video encoding and decoding delay is 4 field time for the frame structure M=1 low delay mode.

4.3 Buffer size

Exact evaluation of necessary buffer size for !multimedia demultiplexing has yet to be carried out.

END

Pack

Syntax elements	No. of bits	Codeword bin, hex, dec	Note
pack_start_code	32	000001BA h	
'0010'	4	0010 b	
system_clock_reference [32..30]	3		
marker_bit	1	1 b	
system_clock_reference [29..15]	15		
marker_bit	1	1 b	
system_clock_reference [14..0]	15		
marker_bit	1	1 b	
marker_bit	1	1 b	
mux_rate	22	12500 d	5 Mbit/s=12500x50bytes/s
marker_bit	1	1b	
system_header			
packet			

System header

Syntax elements	No. of bits	Codeword bin/hex/dec	Note
system_header_start_code	32	000001BB h	
header_length	16	0012 h	18 bytes
marker_bit	1	1 b	
rate_bound	22	12500 d	
marker_bit	1	1 b	
audio_bound	6	000001 b	1 stream of G.722
fixed_flag	1	1 b	fixed bitrate operation
CSPS_flag	1	1 b	"constrained" in this case study; extension to MPEG-2 may need the profile approach
system_audio_lock_flag	1	0 b	16 kHz is not locked with SCR
system_video_lock_flag	1	1 b	13.5 MHz is locked with SCR
marker_bit	1	1 b	
video_bound	5	00001 b	1 stream of H.26X
reserved_byte	8	FF h	
stream_id	8	?? h	G.722 audio, No.0
'11'	2	11 b	
STD_buffer_bound_scale	1	0 b	in units of 128 bytes
STD_buffer_size_bound	13	1 d	128 bytes
stream_id	8	?? h	H.26X, No.0
'11'	2	11 b	
STD_buffer_bound_scale	1	1 b	in units of 1024 bytes
STD_buffer_size_bound	13	128 d	128K bytes, assuming maximum 1 Mbits per picture (???)
stream_id	8	?? h	CCITT defined Data, No.0
'11'	2	11 b	
STD_buffer_bound_scale	1	0 b	in units of 128 bytes
STD_buffer_size_bound	13	1 d	128 bytes
stream_id	8	?? h	CCITT defined Control No.0
'11'	2	11 b	
STD_buffer_bound_scale	1	0 b	in units of 128 bytes
STD_buffer_size_bound	13	1 d	128 bytes

Audio Packet

Syntax elements	No. of bits	Codeword bin/hex/dec	Note
packet_start_code_prefix	24	000001 h	
stream_id	8	?? h	G.722 audio, No.0
packet_length	16	00FF h	7+58=63
stuffing byte	8xn	FF h	n=0; it can be 0-16
'01'	2	01 b	
STD_buffer_scale	1	0 b	in units of 128 bytes
STD_buffer_size	13	1 d	128 bytes
'0010'	4	0010 b	PTS only
presentation_time_stamp [32..30]	3		
marker_bit	1	1 b	
presentation_time_stamp [29..15]	15		
marker_bit	1	1 b	
presentation_time_stamp [14..0]	15		
marker_bit	1	1 b	
packet_data_byte	58x8		G.728 coded audio data

Video Packet

Syntax elements	No. of bits	Codeword bin/hex/dec	Note
packet_start_code_prefix	24	000001 h	
stream_id	8	?? h	H.26X, No.0
packet_length	16	1119 h	16/15 + 4361/4362 = 4377
stuffing byte	8x8 or 7x8	FF h	n=8or 7; it can be 0-16
'01'	2	01 b	
STD_buffer_scale	1	1 b	in units of 1024 bytes
STD_buffer_size	13	64 d	64 Kbytes
'0010'	4	0010 b	PTS only
presentation_time_stamp [32..30]	3		
marker_bit	1	1 b	
presentation_time_stamp [29..15]	15		
marker_bit	1	1 b	
presentation_time_stamp [14..0]	15		
marker_bit	1	1 b	
packet_data_byte	4361x8 or 4362x8		H.26X coded audio data

Data Packet

Syntax elements	No. of bits	Codeword bin/hex/dec	Note
packet_start_code_prefix	24	000001 h	
stream_id	8	?? h	CCITT defined Data, No.0
packet_length	16	0647 h	7+800=807
stuffing byte	0x8	FF h	n=0; it can be 0-16
'01'	2	01 b	
STD_buffer_scale	1	0 b	in units of 128 bytes
STD_buffer_size	13	1 d	128 bytes
'0010'	4	0010 b	PTS only
presentation_time_stamp [32..30]	3		
marker_bit	1	1 b	
presentation_time_stamp [29..15]	15		
marker_bit	1	1 b	
presentation_time_stamp [14..0]	15		
marker_bit	1	1 b	
packet_data_byte	200x8		e.g. JPEG still picture data

Terminal-to-terminal control Packet

Syntax elements	No. of bits	Codeword bin/hex/dec	Note
packet_start_code_prefix	24	000001 h	
stream_id	8	?? h	CCITT defined Control, No.0
packet_length	16	00C7 h	7+200=207
stuffing byte	0x8	FF h	n=0; it can be 0-16
'01'	2	01 b	
STD_buffer_scale	1	0 b	in units of 128 bytes
STD_buffer_size	13	1 d	128 bytes
'0010'	4	0010 b	PTS only
presentation_time_stamp [32..30]	3		
marker_bit	1	1 b	
presentation_time_stamp [29..15]	15		
marker_bit	1	1 b	
presentation_time_stamp [14..0]	15		
marker_bit	1	1 b	
packet_data_byte	200x8		end-to-end control

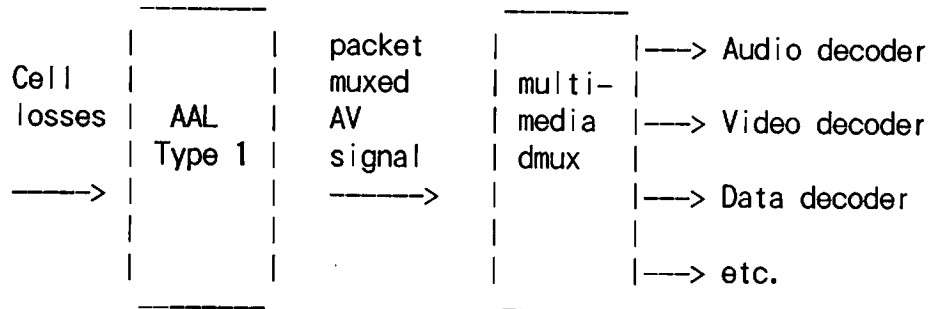
Padding Packet

Syntax elements	No. of bits	Codeword bin/hex/dec	Note
packet_start_code_prefix	24	000001 h	
stream_id	8	BE h	padding
packet_length	16	m+1 d	
'0000 1111'	8	0F h	
'1111 1111'	mx8	FF h	m=0-65534

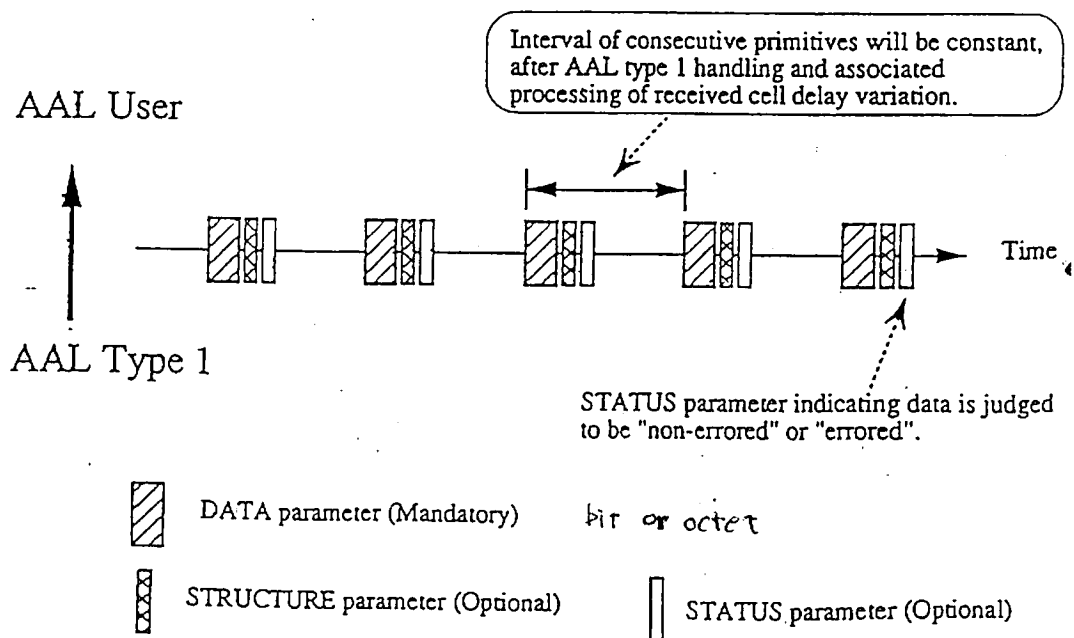
Impacts of cell losses when MPEG-1 Systems is used for multimedia multiplexing

1. Assumption

ISO/IEC 11172-1 (MPEG-1 Systems) is used for multimedia multiplexing and demultiplexing. We assume each packet is filled with a single medium such as video, audio, data, control, etc.. (Annex 1) The model is illustrated as in the following figure;



AAL type 1 is used. When a cell loss occurs, AAL stuffs it with dummy bits and indicates error as shown in the following figure [1].

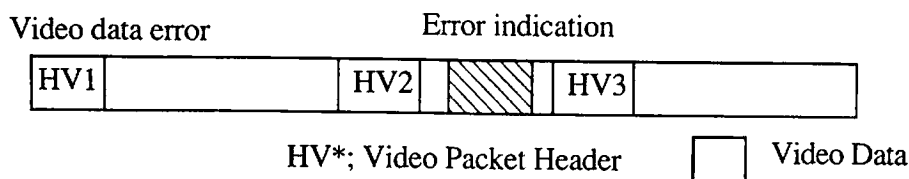


2. Impact caused by cell losses

Following two types of errors can be considered. *1

a. Type 1 error

Condition; Cell losses occur outside packet header (multimedia demultiplexing is correct)



Total number of bits passed on to the decoder does not change, but some of them are erroneous.

Bitstream sent by the video encoder

...V01,V02,V03,V04,V05,V06,V07,V08,V09,V10,V11,V12,V13,V14,...

Bitstream received by the video decoder

...V01,V02,V03,V04,V05,X06,X07,X08,X09,X10,V11,V12,V13,V14,...

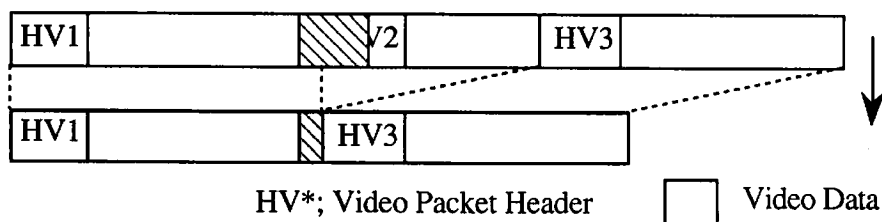
where V_{nn} indicates a group of bits (cell payload or part of it) and X indicates erroneous group of bits.

Impact; There is no difference whether cell losses occur in the multiplexed bitstream or in elementary video bitstream. The probability of cell loss in each elementary bit stream is proportionate to its bit rate compared to the cell loss probability of multiplexed bitstream.

b. Type 2 error

Condition; Cell losses occur at packet header

Packet header is errored, and the packet is lost.



Some group of bits are lost in the forwarded bitstream. This happens when a packet header becomes erroneous and the demultiplexer knows that it is erroneous. Multimedia demultiplexer can not forward this packet to any elementary bitstream decoder since it does not know what medium the packet belongs to.

Bitstream sent by the video encoder

...V01,V02,V03,V04,V05,V06,V07,V08,V09,V10,V11,V12,V13,V14,...

Bitstream received by the video decoder

...V01,V02,V03,V04,V05,V11,V12,V13,V14,...

where V06 through V10 are contained in the lost packet that has been caused by a cell loss or other transmission errors.

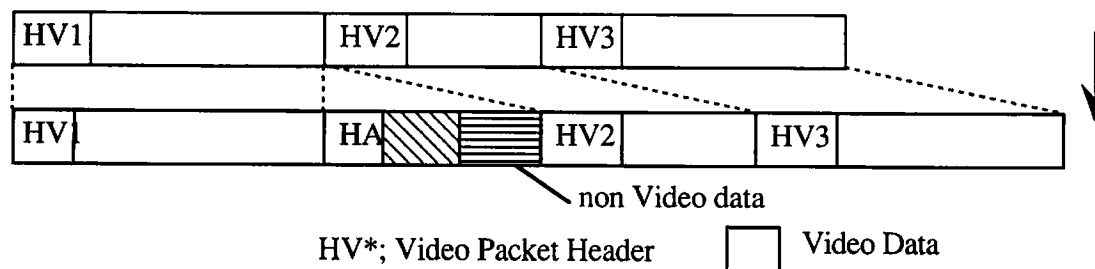
Impact; One packet is lost without error indication to video, audio and other elementary bitstream decoders.

*1; If there is no error indication from AAL to the multimedia demultiplexer, the following type 3 error may occur.

c. Type 3 error

Condition; Cell losses occur at packet header

Packet header is errored, and an irrelevant packet is inserted.



Some group of irrelevant bits are inserted in the forwarded bitstream.

This happens when a packet header becomes erroneous and the packet is forwarded to another elementary bitstream decoder.

Bitstream sent by the video encoder

...V01,V02,V03,V04,V05,V06,V07,V08,V09,V10,V11,V12,V13,V14,...

Bitstream received by the video decoder

...V01,V02,V03,V04,X05,X06,X07,X08,X09,X10,V05,V06,V07,V08,V09,V10,V11,V12,V13,V14,...

where X05 through X10 were misinserted.

Impact; Irrelevant packet is inserted without error indication to video, audio and other elementary bitstream decoders.

3. Consideration

Impact of cell losses in the multiplexed bitstream and impact of cell losses in the elementary video bitstream differ in type 2 and type 3 errors. The probability of type 2 error (packet header loss) is $1/n$ of cell loss probability when the packet interval of a particular medium is n cells. It may sometimes not be negligible. The probability of type 3 error is very low and can be negligible. The implications for type 2 errors are;

- Demultiplexer cannot indicate occurrence of error to the decoder.
- Number of lost bits is usually more than one cell.

In either case, decoder has to search for a next start code. How much importance the error indication from demultiplexer to decoder has should be studied.

4. Conclusion

When MPEG-1 Systems are used for multimedia multiplexing, packet losses without error indication to video and audio decoders occur. The packet loss probability is $1/n$ of cell loss probability when the packet interval of a particular medium is n cells. Once a packet is lost, number of lost bits is usually more than one cell.

The following points should be studied;

- How much importance does error indication from demultiplexer to decoder have?
- Does packet loss, which is equivalent to multiple cell losses, cause a serious problem?

[1] IVS Baseline document, page. 38, June 1992.