INTERNATIONAL TELECOMMUNICATION UNION

CCITT G.764

THE INTERNATIONAL
TELEGRAPH AND TELEPHONE
CONSULTATIVE COMMITTEE

GENERAL ASPECTS OF DIGITAL TRANSMISSION SYSTEMS; TERMINAL EQUIPMENTS

VOICE PACKETIZATION – PACKETIZED VOICE PROTOCOLS

Recommendation G.764

Geneva, 1990

FOREWORD

The CCITT (the International Telegraph and Telephone Consultative Committee) is a permanent organ of the International Telecommunication Union (ITU). CCITT is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The Plenary Assembly of CCITT which meets every four years, establishes the topics for study and approves Recommendations prepared by its Study Groups. The approval of Recommendations by the members of CCITT between Plenary Assemblies is covered by the procedure laid down in CCITT Resolution No. 2 (Melbourne, 1988).

Recommendation G.764 was prepared by Study Group XV and was approved under the Resolution No. 2 procedure on the 14 of December 1990.

CCITT NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication Administration and a recognized private operating agency.

© ITU 1990

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the ITU.

Recommendation G.764

VOICE PACKETIZATION — PACKETIZED VOICE PROTOCOLES

1 Introduction

This Recommendation defines a packet voice protocol for speech packetization in permanent virtual circuit applications. The principal application of the packetized voice protocol (PVP) is at the primary rate and in fractional primary rate applications.

The protocol defines formats and procedures for the transport of voice information and channel associated signalling over a wideband packet network.

The Recommendation accommodates additional future types including optional internetworking capabilities with the digital cellular radio networking applications currently under development. The extension of this Recommendation for baseband facsimile traffic is currently under study.

This Recommendation does not address performance issues.

This Recommendation does not address methods for coding speech samples, although particular recommended coding algorithms are specified in the protocol (e.g. the algorithms in Recommendation G.726). In particular, the Recommendation allows dynamic bandwidth allocation and graceful congestion control when the speech samples are coded with embedded algorithms such as those specified in Recommendation G.727.

This Recommendation does not address the following issues:

1) services based on the interface;

2) implementation techniques;

3) performance guidelines relative to the use of packetized speech;

4) equipment aspects;

5 trunk signalling, link establishment and call establishment procedures for switched virtual circuits;

6) data only issues, combined data/voice issues and frame relay issues;

7) speech packetization in Asynchronous Transfer Mode (ATM) (B-ISDN) systems.

2 Overview

This Recommendation contains the specification of a protocol for packetized speech (PVP). PVP defines formats and procedures for the transport of voice information and channel associated signalling over a packet network.

Before packetization, the input speech samples may be coded at the originating endpoint of the transmitting side by one of the coding methods indicated in this document. The stream of coded speech is transformed into packets with the format specified in this document. The samples are collected over a period of 16 ms and divided into blocks of 128 bits each. Silent intervals may be removed. The blocks are arranged to facilitate block dropping.

Periods of activity and inactivity are respectively called “bursts” and “gaps”. It is not necessary to transmit packets during gaps.

The terminating end at the receiving side reconstructs a continuous stream of speech from the incoming packets using the information in the packet header. The build-out delay procedure described in this document compensates for the variable delay that packets may experience within the network. Packets that arrive before their scheduled play-out time are placed in the proper sequence in a packet queue. Packets that arrive after their scheduled play-out time are discarded. The voice packet header contains information about the level of noise that was measured by the originating endpoint. The terminating endpoint uses this information to play out a matching noise level.

An additional feature of PVP is the ability to drop blocks from a packet as a congestion control mechanism. The *n*-th block consists of the *n*-th bit from each sample collected during the sampling interval. The packet header indicates the number of droppable blocks contained in the packet. Congested nodes may use this information to drop the least significant block from packets to abate the congested state.

The signalling associated with each voice connection shall be transported in signalling packets. Signalling packets shall be sent separately on a different logical channel. Transport of the signalling information requires a set of procedures, similar to those for voice transport, which are described in this document.

Note — In a national network, the time stamp (TS) and the build-out procedures of §§ 5.1.1, 5.2 and 6.3 may be replaced by a fixed-delay for the first packet. At the originating end, the TS will be set to zero. In an intermediate node, the TS field will not be updated. However, the build-out procedure will be used always on network-network and user-network interfaces.

3 Formats

3.1 *Physical layer*

For operations at 1536 kbit/s or 1920 kbit/s, the electrical characteristics and formats of the interface are those defined in Recommendations G.703, G.704 and I.431 for the primary rates of 1544 kbit/s and 2048 kbit/s, respectively. The packetized signal consists of one digital stream sent over conventional primary rate facilities. Hybrid situations containing one or more $N \times 64$ kbit/s packet streams and M conventional 64 kbit/s channels are also considered.

3.1.1 *Bit inversion*

For primary rate applications that require code restrictions that maintain one's density, bit inversion is necessary so that the combined result of bit stuffing and bit inversion is to prevent the all 0 octet and to satisfy the one's density requirements of restricted DS 1 facilities.

3.1.2 *Order of transmission*

Bit 1 is the least significant bit (LSB) and is transmitted first. Bit 8 is the most significant bit (MSB) and is transmitted last.

3.2 *Link layer*

The link layer of PVP uses a similar approach to Recommendation Q.921/I.441 with the additions indicated in this Recommendation. Frames that transport voice and frames that transport channel associated signalling are assigned different layer 2 addresses, i.e. they are carried on two separate logical links. This, together with the use of a different unnumbered frame type for each type of traffic, provides an additional measure of security to protect from the misrouting of signalling information.

3.2.1 *Address field*

The address field is two octets in length with the first bit of each defined as an extension bit and bit 2 of octet 1 defined as the command/response (C/R) bit. The 13 bits that remain are concatenated to form a single data link connection identifier (DLCI). Address assignment starts with 128 and ends with 8063. Layer 2 addresses are already assigned and the implementation starts from the DLCI_ASSIGNED state.

3.2.2 *Command/response bit*

The C/R bit (bit 2 of octet 1) is set to 0.

3.2.3 *Frame types*

The following two frame types are allowed in PVP.

3.2.3.1 *Unnumbered information frames*

When a layer 3 or management entity requests unacknowledged information transfer, the unnumbered information (UI) command is used to send information to its peer without affecting data link layer variables. UI command frames do not carry a sequence number and therefore, the UI frame may be lost without notification.

The control field for the UI command frame is a single octet in length. The format and encoding are the same as specified in Recommendation Q.921/I.441. The UI frame is used to transport channel associated signalling.

3.2.3.2 *Unnumbered information with header check frame*

The unnumbered information with header check (UIH) frame has the same applications as the UI frame. The difference between the two is that the cyclic redundancy check (CRC) sequence is derived over the frame and packet headers (the first 8 octets excluding flags) rather than over the entire frame. The check sequence fills the last two octets of the UIH frame.

The control field of the UIH is a single octet in length and is shown in Figure 1/G.764.

8	7	6	5	4	3	2	1	Bit number
1	1	1	P	1	1	1	1	

FIGURE 1/G.764

Control field of the UIH frame

The UIH frame is used to transport voice (see Note).

Note — The CRC check of the UIH protects 8 octets which contain the address field (to ensure correct delivery), the control field (to guarantee the validity of the frame type) and the layer 3 header. It does not protect the voice information because voice traffic is more sensitive to delay due to retransmission than to bit errors and because this allows reduction of the voice information by block dropping under congestion without recalculating the CRC check. As a consequence, the test for invalid frames in § 3.2.7 uses the minimum frame length of 10 octets.

3.2.4 *Poll bit*

The poll bit (P) is bit 5 of the UI/UIH frame control field. The P bit shall be set to 0.

3.2.5 *Check sequence*

The check sequence (CS) algorithm is the same as that described in ISO Recommendation ISO-3309. The CS field shall be a 16-bit sequence. It shall be the ones complement of the sum (modulo 2) of:

1) The remainder of (x raised to k power) ($x^{15} + x^{14} + x^{13} + x^{12} + x^{11} + x^{10} + x^9 + x^8 + x^7 + x^6 + x^5 + x^4 + x^3 + x^2 + x^1 + 1$) divided (modulo 2) by the generator polynomial $x^{16} + x^{12} + x^5 + 1$, where k is the number of bits in the frame existing between, but not including, the final bit of the opening flag and the first bit of the first octet of the non-droppable octets for the header check sequence (HCS) or the first bit of the check sequence for the frame check sequence (FCS), excluding bits inserted for transparency, and

2) the remainder of the division (modulo 2) by the generator polynomial $x^{16} + x^{12} + x^5 + 1$, of the product of x^{16} by the content of the frame existing between, but not including, the final bit of the opening flag and the first bit of the first octet of the non-droppable octets for the HCS or the first bit of the CS for the FCS, excluding bits inserted for transparency.

As a typical implementation at the transmitter, the initial content of the register of the device computing the remainder of the division is preset to all ones and is then modified by division by the generator polynomial (as described above) of the address, control and appropriate portion of the information fields; the ones complement of the resulting remainder is transmitted as the 16-bit CS.

As a typical implementation at the receiver, the initial content of the register of the device computing the remainder is preset to all ones. The final remainder after multiplication by x^{16} and then division (modulo 2) by the generator polynomial $x^{16} + x^{12} + x^5 + 1$ of the serial incoming protected bits and the CS, will be 0001110100001111 (x^{15} through x^0, respectively) in the absence of transmission errors.

3.2.6 *Frame abort*

Receipt of seven or more contiguous 1 bits shall be interpreted as a frame abort, and the link layer entity shall ignore the frame currently being received. A frame following an abort must begin with an opening flag.

3.2.7 *Invalid UI/UIH frames for PVP*

For the purposes of PVP, an invalid UI/UIH frame is a frame which:

1) is not properly bounded by two flags; or

2) has fewer than 10 octets between flags; or

3) has greater than 490 octets between flags; or

4) does not consist of an integral number of octets prior to zero bit insertion or following zero bit extractions; or

5) contains a FCS error.

Invalid frames shall be discarded without notification to the sender. No action is taken as the result of that frame.

3.3 *Packet layer*

The packet layer procedures apply to the information transfer phase only. Call control procedures are outside the scope of this Recommendation.

3.3.1 *Voice packet format*

The format of voice packets is shown in Figure 2/G.764 within the UIH voice frame.

Note — The reserved bits are set to 0 at the originating endpoint. They shall be ignored at the terminating endpoint. They shall not be used for testing or maintenance purposes in anticipation of possible future uses.

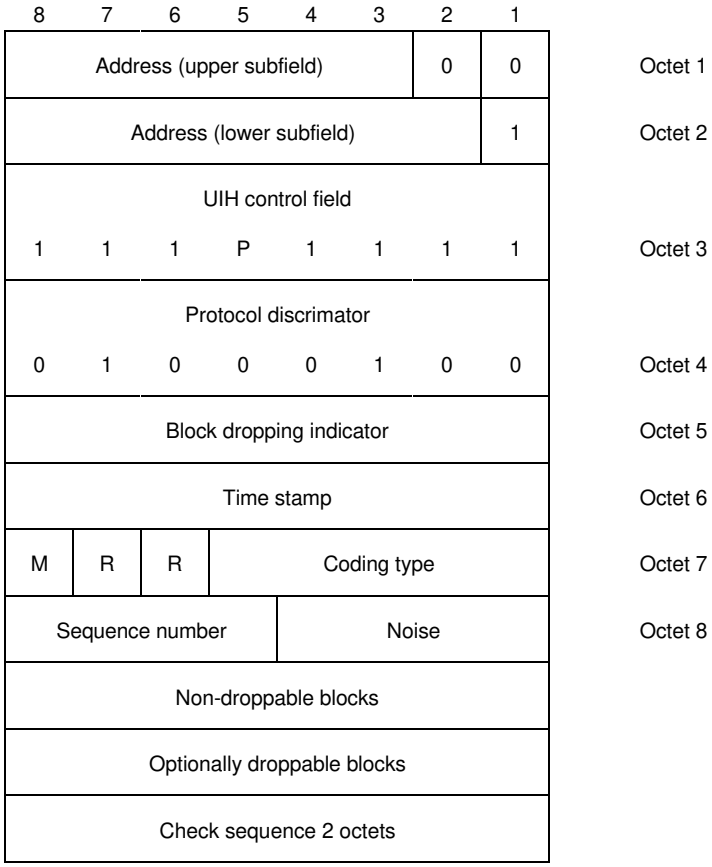

FIGURE 2/G.764

UIH voice frame format

3.3.1.1 *Protocol discriminator*

The protocol discriminator (PD) field is the first octet of the packet header (octet 4, the frame in Figure 2/G.764). Its value for the PVP is given in Figure 3/G.764.

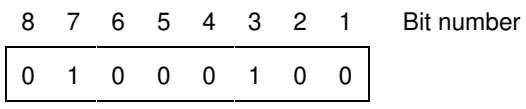

FIGURE 3/G.764

Protocol discrimator for the PVP

3.3.1.2 *Block dropping indicator*

The block dropping indicator (BDI) tracks the status of block dropping within the voice packet. A block consists of bits of the same significance collected from all speech samples that are packetized. The size of the block is 128 bits, which, for a sampling rate of 8 kHz, corresponds to a packetization interval of 16 ms. Blocks are arranged in decreasing order of significance.

The format of the BDI is shown in Figure 4/G.764.

Bit number							
8	7	6	5	4	3	2	1
R	R	M2	M1	R	R	C2	C1

R Reserved for future use

FIGURE 4/G.764

Block dropping indicator format

The combination C1 and C2 form the C-subfield that indicates the number of blocks that are still droppable at any intermediate node in the network, as shown in Table 1/G.764:

TABLE 1/G.764

C-subfield codings

C2	C1	No. of droppable blocks
0	0	0 blocks
0	1	1 block
1	0	2 blocks
1	1	3 blocks

As blocks are dropped from the packet, the value in the C-subfield is decremented to reflect the number of blocks still available for dropping.

The combination M1 and M2 forms the M-subfield that indicates the total number of blocks that can be discarded from the packet as it traverses the network during periods of network congestion, as shown in Table 2/G.764. The value in the M-subfield is not changed from its initial value.

TABLE 2/G.764

M-subfield codings

M2	M1	No. of droppable blocks
0	0	0 blocks
0	1	1 block
1	0	2 blocks
1	1	3 blocks

For fixed rate coding, both the M-subfield and the C-subfield are set to 0.

3.3.1.3 *Time stamp*

The TS records the cumulative variable queueing delays experienced by a packet in traversing the network with a resolution of 1 ms. To prevent wrap-around, the maximum valid value in the TS field shall not exceed 200 ms. If, after update, the variable delay exceeds 20 ms, the value is set to 20 ms.

3.3.1.4 *Coding type*

The coding type (CT) field indicates the method of coding the speech samples at the originating endpoint before packetization. Figure 5/G.764 shows the valid encodings for the field.

When the coding type for voice-band signals is fixed or embedded adaptive differential pulse-code modulation (ADPCM), the polarity of ADPCM samples complies with Recommendations G.726/G.727[1)]. When the coding type for voice-band signals is 8-bit pulse code modulation (PCM), the polarity of PCM samples complies with Recommendation G.711.

3.3.1.5 *M-bit*

The M-bit is set to 1 for all packets except for the last packet of a burst where it is set to 0. The M-bit may be used by the terminating end point to recover from packet loss.

3.3.1.6 *Sequence number*

The sequence number (SEQ) is used by endpoints in the build-out process to:

1) determine the first packet of a burst; and

2) determine if a packet has been lost.

The SEQ, in conjunction with the TS, allows the removal of variability in network delay.

SEQ 0 is placed in the first packet of a voice burst. Subsequent packets in the same burst are given numbers 1 to 15, rolling back to 1.

3.3.1.7 *Noise field*

The noise field indicates a background noise level as shown in Figure 6/G.764. The receiving end uses this field to determine the level of the background noise that may be played in the absence of packets.

---

1) Recommendation G.726 completely replaces Recommendations G.721 and G.723 published in Volume III.4 of the Blue Book. It should be noted that systems designed in accordance with Recommendation G.726 will be compatible with systems designed in accordance with the Blue Book versions.

Bit number					Coding type[a]
5	4	3	2	1	
0	0	0	0	0	8 bits/sample[b]
0	0	0	0	1	1 bit/sample[b]
0	0	0	1	0	2 bits/sample[b]
0	0	0	1	1	3 bits/sample[b]
0	0	1	0	0	4 bits/sample[b]
0	0	1	0	1	5 bits/sample[b]
0	0	1	1	0	6 bits/sample[b]
0	0	1	1	1	7 bits/sample[b]
0	1	0	0	0	8 bit A-law PCM[c]
0	1	0	0	1	8 bit μ-law PCM[c]
0	1	0	1	0	2 bits/sample ADPCM[d]
0	1	0	1	1	3 bits/sample ADPCM[d]
0	1	1	0	0	4 bits/sample ADPCM[d]
0	1	1	0	1	5 bits/sample ADPCM[d]
0	1	1	1	0	Reserved for future use
0	1	1	1	1	Reserved for future use
1	0	0	0	0	Reserved for future use
1	0	0	0	1	Reserved for future use
1	0	0	1	0	Reserved for future use
1	0	0	1	1	Reserved for future use
1	0	1	0	0	(4,2) embedded ADPCM[e]
1	0	1	0	1	(5,2) embedded ADPCM[e]
1	0	1	1	0	Reserved for future use
1	0	1	1	1	Reserved for future use
1	1	0	0	0	(8,6) embedded ADPCM[f]
1	1	0	0	1	Reserved for future use
		...			
		...			
1	1	1	1	1	Reserved for future use

a) Reserved coding types will be assigned as other voice coding algorithms are standardized. In particular, coding types will be assigned for 8 kbit/s and 4 kbit/s coding algorithms intended for digital cellular radio applications. The specific voice packet information field format and length, and the voice transport procedures, will depend on the coding algorithm. In particular, these matters will depend on whether the coding algorithm has a natural block length, has variable rates, or has droppable bits.

b) This is to carry 8 × N kbit/s on transparent channels.

c) Interoperability issues is an item for further study.

d) Recommendation G.726.

e) Applications of the (5,4), (5,3), (4,3), (3,3), and (3,2) ADPCM algorithms are left for study.

f) Recommendation G.722. The procedure for going from an H.221 frame structure to Recommendation G.764 is an item for further study.

FIGURE 5/G.764

Coding type (CT) format

Bit number	Noise leve
4321	(dBmc0)
0000	Idle Code
0001	16.6
0010	19.7
0011	22.6
0100	24.9
0101	26.9
0110	29.0
0111	31.0
1000	32.8
1001	34.6
1010	36.2
1011	37.9
1100	39.7
1101	41.6
1110	43.8
1111	46.6

FIGURE 6/G.764

Noise field format

3.3.1.8 *Voice information field*

The voice information field contains blocks arranged according to the significance of the bits. The first block contains the MSBs of all samples, the second contains the second MSBs and so on. Within a block, bits are ordered according to their corresponding sample numbers.

Figure 7/G.764 shows the bit nomenclature convention before packetization and the information field format after packetization. Figure 8/G.764 depicts the format of the entire voice packet where the voice is coded using an embedded (5,2) ADPCM algorithm. Here, up to 3 blocks can be dropped. Note the most significant bit for PCM is the sign bit.

3.3.2 *Signalling packet format*

The format of the channel associated signalling packets is shown in Figure 9/G.764 within a UI signalling frame format. See § 3.3.1 regarding the setting of the reserved bits.

3.3.2.1 *Protocol discriminator*

The format and encoding of the PD field are the same as in the voice packet format (see § 3.3.1.1).

3.3.2.2 *Block dropping indicator*

The format of the BDI field is the same as in the voice packet format (see § 3.3.1.2). Both the C-subfield and the M-subfield are set to 0.

Sample number				
S1	MSB/S1	(MSB1-1)/S1	· · ·	LSB/S1
S2	MSB/S2	(MSB-1)/S2	· · ·	LSB/S2
.			.	
.			.	
.			.	
S128	MSB/S128	(MSB-1)/S128	· · ·	LSB/S128

a) Pre-packetization bit format

Bit number	8	7		1
MSB block	MSB/S8	MSB/S7	· · ·	MSB/S1
			. .	
	MSB/S128	MSB/S127	· · ·	MSB/S121
MSB-1 block	(MSB-1)/S8	(MSB-1)/S7	· · ·	(MSB-1)/S1
			. . .	
	(MSB-1)/S128	(MSB-1)/S127	· · ·	(MSB-1)/S121
. . .			. . .	
LSB block	LSB/S8	LSB/S7	· · ·	LSB/S1
			. . .	
	LSB/S128	LSB/S127	· · ·	LSB/S121

b) Voice packet information field

FIGURE 7/G.764

Bit-ordering within voice packet information fields

Frame header	Layer 3 header	Most significant non-droppable bits	2nd Most significant non-droppable bits	Last-128-droppable bits	Second-128-droppable bits	First-128-droppable bits	Frame trailer

FIGURE 8/G.764

Block arrangement for (5,2) ADPCM coding

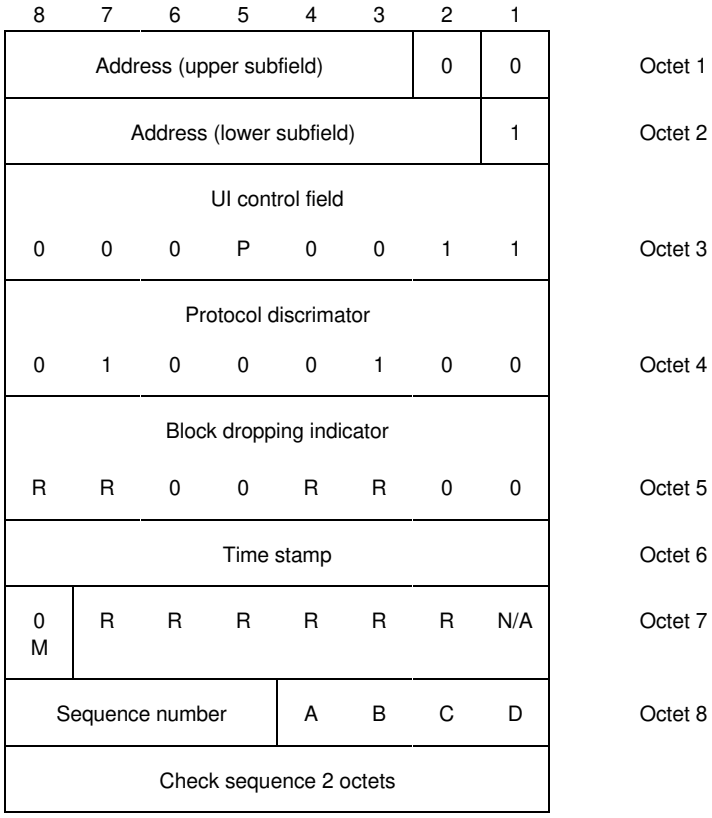

FIGURE 9/G.764

UIH signalling frame format

3.3.2.3 *Time stamp*

The format of the TS field is the same as in the voice packet format (see § 3.3.1.3).

3.3.2.4 *Normal/alarm state bit*

The normal/alarm (N/A) bit is used to transfer information on alarm status across a virtual circuit in the direction of transmission from the full rate access side to the packet side. The N/A bit set to 0 indicates normal operation. The N/A bit set to 1 indicates the existence of an alarm on the full rate access facility or error condition on the virtual circuit.

3.3.2.5 *M-bit*

The M-bit shall be set to 0 in all signalling packets.

3.3.2.6 *Sequence number*

The SEQ for signalling packets is always set to 0.

3.3.2.7 *ABCD signalling bits*

The originating endpoint uses the ABCD signalling bits to indicate to the terminating endpoint on the far side the current signalling state of the full rate access channel in the direction of transmission. The value of the A bit alone is meaningful in two-state signalling systems. The value of the A and B bits alone are meaningful in four-state signalling systems. The values of all ABCD bits are meaningful in 16-state signalling systems. The values of the ABCD bits have no significance when there is no channel associated signalling.

The number of signalling states must be the same for both the originating and terminating endpoints on a virtual circuit basis. The values coded in this field depend on the framing type and the number of signalling states.

4 Link layer procedures

4.1 *Addressing*

Voice and signalling transport packets are transmitted on different layer 2 addresses.

4.2 *Endpoint procedures*

These procedures apply for UI and UIH frames.

4.2.1 *Transmitting UI frames*

Information received by the link layer entity from layer 3 by means of a DL-UNIT-DATA request shall be transmitted as unnumbered information with a FCS. The P bit shall be set to 0. The list of all primitives used in the procedures is given in § 9.

4.2.2 *Transmitting UIH frames*

Information received by the link layer entity from layer 3 by means of a DL-UNIT-H-DATA request shall be transmitted as unnumbered information with an HCS. The P bit shall be set to 0.

4.2.3 *Receiving UI frames*

When a link layer entity is not in a receiver busy condition and receives a valid UI frame, the link layer entity shall pass the information field of this frame to layer 3 using the primitive DL-UNIT-DATA indication.

4.2.4 *Receiving UIH frames*

When a link layer entity is not in a receiver busy condition and receives a valid UIH frame, the link layer entity shall pass the information field of this frame to layer 3 using the primitive DL-UNIT-H-DATA indication.

4.3 *Intermediate node procedures*

4.3.1 *Transmitting a frame*

Whenever a frame is received from the link layer entity receive procedure, the frame shall be transmitted with the same frame type, including the P bit value, and the C/R bit value as in the received frame.

4.3.2 *Receiving a frame*

Detected invalid frames (e.g. failed FCS or HCS, unassigned DLCI) shall be discarded with no indication passed to layer 3. The control field of the frame shall be examined. Upon recognition of UI or UIH frame types, the frame is passed to layer 3 with the DL-PVP-DATA indication primitive for UI frames and DL-PVP-H-DATA indication primitive for UIH frames.

The Address and CS are the only fields modified during the link layer procedure.

5 Voice transport procedures

Voice transport procedures are divided into originating, intermediate and terminating node (endpoint) procedures. Originating endpoints are nodes where user data are formulated into PVP packets for transport. Intermediate nodes are nodes that do not alter the packet format, but simply receive and transport PVP packets. Terminating endpoints are the destination nodes for PVP packets. It is assumed that the processing of primitives requires a fixed amount of time. Any time variance in the processing of primitives shall be accounted for in the value of the timer TVDELAY_V.

Figure 10/G.764 shows a functional viewpoint of an endpoint node, which shows that it consists of an originating endpoint and a terminating endpoint.

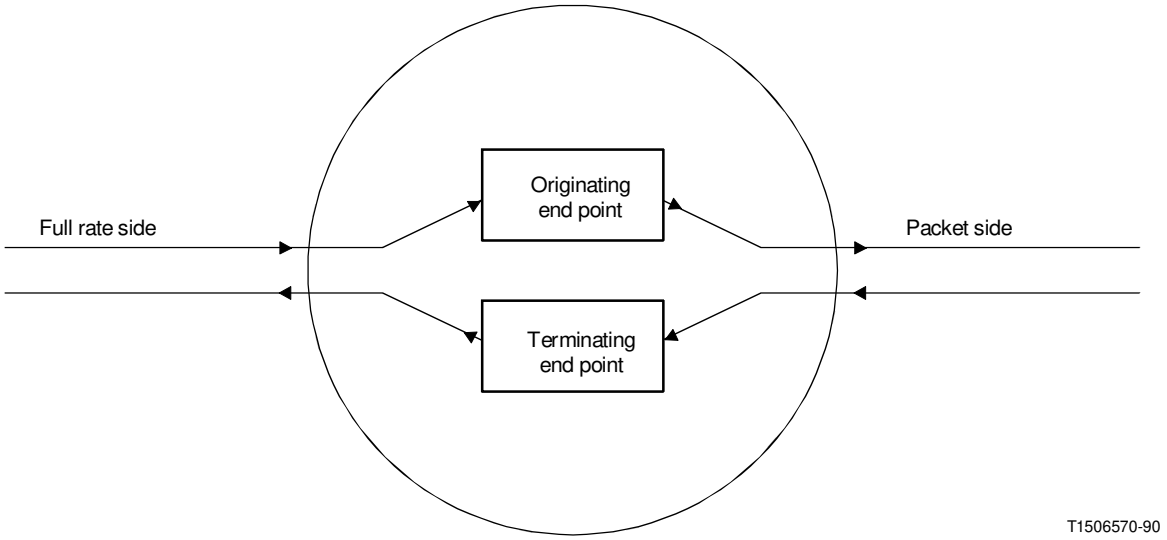

FIGURE 10/G.764

Endpoint node

5.1 *Originating endpoint procedures*

The originating endpoint receives segmented data from a higher layer entity via the primitives PL-START request(CODE, NOISE), PL-DATA request(CODE, NOISE) and PL-STOP request(CODE, NOISE). These primitives include information on the type of encoding and noise level associated with the packet.

5.1.1 *Receipt of PL-START request primitive*

The higher level entity will send to the layer 3 entity, the PL-START request(CODE, NOISE) primitive after it has collected all the samples of the first packet. When the originating endpoint receives the PL-START request(CODE, NOISE) primitive, the layer 3 entity shall start the timer TVDELAY_V associated with the first packet and shall form a voice packet with M-bit set to 1 and SEQ set to 0. The BDI, CT and Noise fields are encoded based on the coding type and noise level indicated in the primitive PL-START request(CODE, NOISE).

The layer 3 entity sets the send sequence state variable (SSEQ) to 1 and checks its own congestion level indicator (CLI) to determine whether blocks should be dropped from the packet and the number of blocks to be dropped. If the CLI is greater than 0, then the block dropping procedures of § 5.4 shall be followed. If the CLI is 0, then the block dropping procedures are omitted. The CLI is a local parameter of the node.

The layer 3 entity shall buffer the packet until notified by the layer 2 entity that layer 1 is ready to transport data. In the absence of facility alarms, this notification shall be conveyed by the primitive DL-L1-READY indication. Upon receipt of the primitive, the timer TVDELAY_V is stopped and its value shall be copied to the TS field. The value of the TS field shall not exceed 200.

The packet is then delivered to the layer 2 entity with the primitive DL-UNIT-H-DATA request.

5.1.2 *Receipt of the PL-DATA request primitive*

After receiving the PL-DATA request(CODE, NOISE) primitive from a higher layer entity, the PVP layer 3 entity shall start the timer TVDELAY_V and form a voice packet with the M-bit set to 1 and the SEQ set to the value of the SSEQ. The BDI, CT and noise fields are encoded on the basis of the corresponding information in the PL-DATA request primitive. The layer 3 entity increments the SSEQ (value from 1 to 15 with a roll over to 1). It shall check the CLI to determine the need for block dropping. If the CLI is greater than 0, then the block dropping procedure of 5.4 shall be followed. If the CLI is 0, then the block dropping procedure shall be omitted.

The layer 3 entity shall await the arrival of the DL-L1-READY indication primitive from layer 2. Upon receipt of the primitive, the timer TVDELAY_V is stopped and its value shall be copied to the TS field. The layer 3 entity shall pass the voice packet to the layer 2 entity for transport using the DL-UNIT-H-DATA request primitive.

5.1.3 *Receipt of the PL-STOP request primitive*

When a higher layer entity detects a gap in the speech, it will continue the packetization until all 128 samples have been packetized. It will then send the PL-STOP request primitive to the layer 3 entity. The layer 3 entity shall follow the procedures outlined above following receipt of the PL-DATA request except that it shall set the M-bit to 0.

5.1.4 *Number of blocks and packetization interval*

The packetization interval is 16 ms. The number of blocks of 128 bits collected during this interval depends on the coding type, as shown in Table 3/G.764.

TABLE 3/G.764

Blocks collected during the packetization interval for various coding types

Coding type	Number of blocks
8 bits/sample	8
1 bit/sample	1
2 bits/sample	2
3 bits/sample	3
4 bits/sample	4
5 bits/sample	5
6 bits/sample	6
7 bits/sample	7
8 bit PCM (A-law or μ-law)	8
2 bits/sample ADPCM	2
3 bits/sample ADPCM	3
4 bits/sample ADPCM	4
5 bits/sample ADPCM	5
(4,2) embedded ADPCM	4
(5,2) embedded ADPCM	5
(8,6) embedded ADPCM	8

5.1.5 *Coder reset*

When the coding type represents that of Recommendations G.722, G.726 or G.727, a voice packet at the beginning of a speech burst (i.e. with SEQ = 0) must start with a reset coder.

Interoperability issues with PCM coding is an issue for further study.

5.2 *Intermediate node procedures*

Upon receipt of the DL-PVP-H-DATA indication primitive, the layer 3 entity will start timer TVDELAY_V associated with that packet. The layer 3 entity shall examine the value encoded in the PD field. If this value matches that for PVP, the layer 3 entity shall examine the CLI which is a system variable. The CLI shall be set by the management entity to indicate the number of blocks to be dropped from packets containing droppable blocks (0, 1, 2 or 3 blocks may be specified).

If the CLI is greater than 0, then blocks may be dropped from the packet according to the block dropping procedures described below. If the CLI is 0, then no block dropping shall occur.

The packet shall then be buffered until the layer 3 entity receives the primitive DL-L1-READY indication from layer 2. Upon receipt of this primitive, the variable delay timer TVDELAY_V shall be stopped and its value shall be used to update the packet's time stamp. The resolution of TVDELAY_V is 1 ms. The value of the TS field shall not exceed 200 ms.

The layer 3 entity shall then pass the information to the layer 2 entity via the DL-PVP-H-DATA request primitive.

5.3 *Terminating endpoint procedures*

Upon receipt of the DL-UNIT-H-DATA indication primitive, the layer 3 entity shall examine the value encoded in the PD field. If this value matches that for PVP, the layer 3 entity will proceed as below.

5.3.1 *Illegal BDI/coding type combinations*

The packet is dropped if the combination of the CT field and the BDI is illegal. The state variable RSEQ is not updated.

Illegal combinations are those where the C-subfield value and/or the M-subfield value of the BDI field exceed the number of droppable bits specified by the coding type. Valid combinations are defined in Table 4/G.764.

TABLE 4/G.764

Valid BDI/CT combinations

Coding type	M-subfield value	C-subfield value
PCM	0	0
Fixed ADPCM	0	0
(m,n) embedded ADPCM	(m-n)	≤ (m-n)

5.3.2 *Wrong packet length*

A voice packet shall be dropped if its length is not consistent with the BDI and CT fields. The state variable RSEQ shall not be updated. The following equation gives the valid packet length based on the BDI subfield values and the CT:

$$l = \frac{[S-(M-C)]\times R\times T}{8} + 5$$

where

l is the packet length in octets

S is the number of bits per sample (from coding type)

M is the value of M-subfield (from BDI)

C is the value of C-subfield (from BDI)

R is the sampling rate (8000 samples/sec)

T is the sampling period (16 ms)

5.3.3 *Play-out procedures*

5.3.3.1 *Decoder reset*

When the coding type represents that of G.722, G.726 or G.727, a voice packet at the beginning of a speech burst (i.e. with SEQ = 0) must cause the decoder to reset. This is done by sending to the management entity the primitive MPL-DECODER-RESET request.

Interoperability issues with PCM coding is an issue for further study.

5.3.3.2 *Build-out delay procedures*

The build-out delay is a system variable that defines at each end the maximum allowable variable delay in the transmission path. The purpose is to mask the variability in the delay that each packet may experience so that all packets are played at regular intervals thereby achieving packet voice synchronization. This value shall be always $\leq$ 199 ms. Packets that experience delays beyond the build-out value are discarded. The policy of packet replacement is left for further study.

Packets that are played out based on the TS value are:

1) packets with SEQ = 0, i.e. the first packet in a voice burst and all signalling packets;

2) the voice packet that follows a missing packet, i.e. SEQ in the received packet is different from receive sequence state variable (RSEQ).

When the time for playing out a packet is based on the TS, the duration it is held before play-out is given by:

(Build-out delay) — (TS value) = (ms to hold packet before play-out).

The packet shall be played out at the end of this period.

Packets with non-zero sequence numbers that are received in sequence are played out without a gap after the preceding packet.

The state variable RSEQ is updated after a voice packet has been scheduled for play-out.

5.3.3.3 *Embedded ADPCM*

For embedded ADPCM, the receiving end determines the algorithm to use for decoding the speech through the BDI and CT fields.

5.3.3.4 *Absence of packets*

In the absence of packets to be played, the M-bit of the previous packet can be used to determine whether an interpolation procedure is necessary.

The M_LAST system variable stores the value of the M-bit in the last packet.

If M_LAST = 0, the gap is legitimate and the terminating endpoint shall play out the background noise level that corresponds to the value encoded in the noise field of the last received voice packet, as given in the table of Figure 6/G.764. If M_LAST = 1, then a packet was lost. Recommended interpolation procedures, e.g. noise fill or last packet replay, are left for further study.

5.4 *Block dropping procedures*

Embedded coding algorithms allow for droppable and non-droppable bits within the packet. Dropping the first droppable bit of each sample corresponds to dropping the last block in the packet.

When the CLI specifies the dropping of 1 or more blocks, the layer 3 entity shall determine from the C-subfield (in the BDI field) of a packet the number of droppable blocks available for dropping. The number of blocks that can be still dropped is given by:

min(value in C-subfield, value in CLI)

The C-subfield of the BDI shall be updated to indicate the number of blocks that can be dropped at the following nodes. This number can be set to 0 if no blocks can be dropped at the following nodes.

6 Signalling transport procedures

6.1 *General principles*

Channel associated signalling is transported across the packet network using signalling packets. To minimize the incorrect delivery of signalling information, channel associated signalling is transported in a UI frame with a different logical address than that of the corresponding UIH frame that transports the voice information.

There are two types of signalling packets: signalling transition packets and refresh packets. Both have the same structure shown in Figure 9/G.764 and are enclosed in UI frames. The originating end sends a signalling transition packet whenever the signalling state changes. It sends refresh packets on a periodic basis to indicate that the link is still active.

The endpoints will generate and receive signalling packets for each virtual circuit provisioned for channel associated signalling. To account for a variety of signalling schemes, the endpoints shall provide for the following variations:

1) No signalling packets.

2) Signalling refresh packets only.

3) Signalling refresh packets and signalling transition packets for 2-state signalling using the A bit. Changes in the A bit result in transition packets while other bits are ignored.

4) Four-state signalling using the A and B bits so that signalling refresh packets and signalling transition packets for changes in the A and B bits result in transition packets while other bits are ignored.

5) Signalling refresh packets and signalling transition packets for 16-state signalling using the A, B, C, and D bits so that transitions in any of the four signalling bits trigger transition packets.

The number of signalling states must be the same for both the originating and terminating endpoints on a virtual circuit basis. The values coded in this field depend on the framing type and the number of signalling states.

It is assumed that the processing of primitives requires a fixed amount of time. Any time variance in the processing of primitives shall be accounted for in the value of the timer TVDELAY_SIG.

6.2 *Originating endpoint procedures*

The management entity of the originating endpoint shall perform the following procedures for each virtual circuit provisioned to support channel associated signalling:

1) Determine, once per extended superframe, the current state of the ABCD bits and determine whether a transition has occurred in accordance with the number of signalling states supported.

2) When a transition occurs, the management entity shall send MPL-SIG-PKT request (A,B,C,D,N/A) primitive to the originating end of the PVP entity. This originating endpoint must then transmit a transition signalling packet containing the current signalling and alarm states.

The originating endpoint starts in the NORM state. In this state, signalling packets (refresh packets) with the current signalling and alarm states as indicated by the N/A bit must be sent at least every TSIG_REF s. The default value for the refresh timer, TSIG_REF, is 10 seconds.

The N/A bit is set to 0 as long as there are no facility alarms (out-of-frame, red or yellow alarms) and that TSIG_KA timer has not expired. The N/A bit shall be set to 1 if there is a facility alarm or if the TSIG_KA timer of the associated terminating end has expired (i.e. the terminating end is in state L_ALARM). Upon the occurrence of a facility alarm, the originating end of the PVP entity shall move from the NORM state to the ALARM state and shall stop transmitting transition packets. It shall continue to send refresh packets every TSIG_REF seconds with the N/A bit set to 1.

The signalling state is frozen until the alarm condition is terminated and the originating endpoint moves back to the NORM state. In the NORM state, the originating endpoint sends packets with the N/A bit set to 0 after the associated terminating endpoint has received a signalling packet.

6.3 *Intermediate node signalling procedures*

Upon receipt of the DL-PVP-DATA indication primitive, the layer 3 entity will start timer TVDELAY_SIG associated with that packet. The layer 3 entity shall examine the value encoded in the PD field. If this value matches that for PVP, the layer 3 entity shall buffer the packet until it receives the primitive DL-L1-READY indication from layer 2. Upon receipt of this primitive, the variable delay timer TVDELAY_SIG shall be stopped and its value shall be used to update the packet's time stamp. The resolution of TVDELAY_SIG is 1 ms. The value of the TS field shall not exceed 200 ms.

The layer 3 entity shall then pass the signalling information to the layer 2 entity via the DL-PVP-DATA request primitive.

6.4 *Terminating endpoint procedures*

The terminating endpoint shall perform the following procedures for each 64 kbit/s stream provisioned for channel associated signalling:

1) In the NORM state: Upon receipt of a signalling packet, the terminating endpoint shall build out the packet according to the time stamp and reinsert the A B C D bits into the PCM bit stream. The build-out procedure is the same as for voice packets (see § 6.3.3.2) except that at play-out time, the layer 3 entity informs the Management Entity of the signalling packet with the primitive MPL-SIG-PKT indication (A,B,C,D,N/A). The terminating endpoint shall continue to use the most recent signalling bits until another signalling packet is received.

2) If no signalling packet is received for a time TSIG_KA (i.e. TSIG_KA has expired), the terminating endpoint shall move to the L_ALARM state and shall initiate trunk conditioning towards the full rate access side. The default value for TSIG_KA is 25 seconds. Upon receipt of a signalling packet with the N/A bit set to 0, the terminating end of the PVP entity shall move to the NORM state and shall terminate the trunk conditioning.

3) In the NORM state: Upon receipt of a signalling packet with the N/A bit set to 1, the terminating end of the PVP entity shall move to the R_ALARM state and shall initiate trunk conditioning towards the full rate access side. It shall return to the NORM state and terminate the trunk conditioning when it receives a signalling packet with the N/A bit set to 0.

4) If the TSIG_KA expires in the R_ALARM state, the terminating end shall move to the L_ALARM state.

5) If a signalling packet arrives with N/A = 1 in the L_ALARM state, the terminating end shall move to the R_ALARM state.

6.5 *Signalling states*

6.5.1 *Originating end signalling states*

The originating end has two signalling states, as follows:

6.5.1.1 *Normal state*

In this state there are no access facility alarms on the full rate access side.

6.5.1.2 *Alarm state*

This is the state when there is a facility alarm on the full rate access side.

6.5.2 *Terminating end signalling states*

The terminating endpoint has three signalling states:

6.5.2.1 *Normal state (NORM)*

The terminating endpoint remains in this state as long as signalling packets arrive with N/A bit set to 0 and TSIG_KA has not expired.

6.5.2.2 *Loss of keep alive alarm (L_ALARM)*

The timer TSIG_KA has expired without the reception of a signalling packet. This indicates that a failure has occurred with the packet portion of the connection and normal signalling packet transport has been interrupted.

6.5.2.3 *Remote alarm (R_ALARM)*

Receipt of signalling packet with N/A bit set to 1 indicates that the far end is experiencing an alarm condition.

7 System variables

7.1 *Send sequence state variable*

Each transmitting endpoint shall have an associated SSEQ that stores the value of the SEQ of the next packet to be transmitted. SSEQ can take on the value of 0 through 15 and is incremented by 1 after each successful packet transmission. When SSEQ has the value 15 and another packet of the same voice burst is transmitted, the value of SSEQ is updated to 1.

7.2 *Receive sequence state variable*

Each terminating endpoint shall have an associated RSEQ that stores the sequence number of the next in-sequence voice packet expected to arrive. RSEQ can take on the value of 0 through 15 and is incremented by 1 (with roll-over to 1) after the voice packet is scheduled for play-out. RSEQ takes on the value 0 only when the last packet of a voice burst has been scheduled for play-out.

7.3 *M_LAST variable*

M_LAST stores the value of the M-bit in the last packet.

7.4 *Congestion level indicator (CLI) variable*

The CLI is set by the management entity to indicate the number of blocks to be dropped from packets containing droppable blocks (0, 1, 2, or 3 blocks may be specified).

8 Protocol parameters

8.1 *Build-out delay*

The value of the build-out delay shall be equal to the maximum allowable variable delay in the transmission path of voice-band traffic. This value shall be $\leq$ 199 ms. The resolution of the build-out delay is 1 ms.

8.2 *TSIG_REF*

The interval between successive transmissions of refresh signalling packets from the originating end point of the node containing the PVP entity. It may take the values of 1, 5, 10 or 20 s. The default value is 10 s.

8.3 *TSIG_KA*

This is the maximum time allowed without receiving a signalling packet at the terminating end of a node containing the PVP entity before it must take recovery actions. TSIG_KA is a multiple of TSIG_REF. The multiplier may be set to 1.5, 2.5, 3.5 or 4.5. The default value of the multiplier is 2.5.

8.4 *TVDELAY_V*

This timer is used to measure the variable queueing delay in a node that a voiceband packet encounters. It is used to update the TS field of a voice-band packet.

8.5 *TVDELAY_SIG*

This timer is used to measure the variable queueing delay in a node that a signalling packet encounters.

It is used to update the TS field of a signalling packet.

9 Summary of primitives

9.1 *Primitives for the interface with layer 2*

The layer 2 primitives used for PVP are described below:

9.1.1 *DL-L1-READY indication*

This primitive is used to indicate to layer 3 that layer 1 is ready for transmission.

9.1.2 *DL-UNIT-DATA request*

The DL-UNIT-DATA request primitive is used by the layer 3 entity to request the transmission of layer 3 messages which are to be transmitted by the data link layer in UI frames using the unacknowledged information transfer service.

9.1.3 *DL-UNIT-DATA indication*

The DL-UNIT-DATA indication primitive is used to indicate receipt by the data link layer of PVP at the terminating endpoint of layer 3 messages that are enclosed UI frames.

9.1.4 *DL-UNIT-H_DATA request*

The DL-UNIT-H-DATA request primitive is used by the layer 3 entity to request the transmission of layer 3 messages which are to be transmitted by the data link layer in UIH frames using the unacknowledged information transfer service.

9.1.5 *DL-UNIT-H-DATA indication*

The DL-UNIT-H-DATA indication primitive is used to indicate receipt by the data link layer of PVP at the terminating endpoint of layer 3 messages that are enclosed in UIH frames.

9.1.6 *DL-PVP-H-DATA indication*

The DL-PVP-H-DATA indication primitive is used to indicate receipt by the data link layer of PVP at intermediate nodes of layer 3 messages in UIH frames.

9.1.7 *DL-PVP-DATA indication*

The DL-PVP-DATA indication primitive is used by the data link layer of PVP intermediate nodes to indicate receipt of layer 3 messages in UI frames.

9.1.8 *DL-PVP-H-DATA request*

The DL-PVP-H-DATA request primitive is used by the layer 3 entity of an intermediate node to indicate to the data link layer a layer 3 message to be transported in a UIH frame.

9.1.9 *DL-PVP-DATA request*

The DL-PVP-DATA request primitive is used by the layer 3 entity of an intermediate node to indicate to the data link layer a layer 3 message to be transported in a UI frame.

9.2 *Primitives for the interface with upper layers*

9.2.1 *PL-START request(CODE, NOISE)*

The PL-START request(CODE, NOISE) primitive is used by the higher layer of the originating endpoint to request that the layer 3 entity begin formatting voice packets with the CT = CODE and the Noise = NOISE.

9.2.2 *PL-DATA request(CODE, NOISE)*

The PL-DATA request(CODE, NOISE) primitive is used by the higher layer of the originating endpoint to continue the formatting voice packets with the CT = CODE and the Noise = NOISE.

9.2.3 *PL-STOP request*

The PL-STOP request primitive is used by upper layers to indicate the end of a speech burst.

9.3 *Primitives for the interface with the management entity*

9.3.1 *MPL-DECODER-RESET request*

The MPL-DECODER-RESET request primitive is used by layer 3 of the PVP originating endpoint to request resetting of the decoder by the management entity.

9.3.2 *MPL-SIG-PKT request(A,B,C,D,N/A)*

This primitive is used by the management entity to request transmission of a transition signalling packet by the PVP layer 3.

9.3.3 *MPL-SIG-PKT indication(A,B,C,D,N/A)*

This primitive is used by the layer 3 entity to inform the management entity of the play-out of a signalling packet by the PVP layer 3.