International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

V.152
(01/2005)

SERIES V: DATA COMMUNICATION OVER THE TELEPHONE NETWORK

Interworking with other networks

Procedures for supporting voice-band data over IP networks

ITU-T Recommendation V.152

ITU-T V-SERIES RECOMMENDATIONS

DATA COMMUNICATION OVER THE TELEPHONE NETWORK

General	V.1–V.9
Interfaces and voiceband modems	V.10–V.34
Wideband modems	V.35–V.39
Error control	V.40–V.49
Transmission quality and maintenance	V.50–V.59
Simultaneous transmission of data and other signals	V.60–V.99
Interworking with other networks	**V.100–V.199**
Interface layer specifications for data communication	V.200–V.249
Control procedures	V.250–V.299
Modems on digital circuits	V.300–V.399

For further details, please refer to the list of ITU-T Recommendations.

ITU-T Recommendation V.152

Procedures for supporting voice-band data over IP networks

Summary

Voice-band data traffic has traditionally been transported by circuit switched systems and equipment. With the advent of the networks optimized for the transport of Internet Protocol (IP), and as a result of its considerable growth and pervasive nature, more and more voice-band data traffic is expected to be carried over IP networks.

Given that voice and voice-band data services remain a significant part of telecommunications, there is a need to ensure a high quality of service for voice and voice-band data carried in part, or wholly, via IP. This Recommendation defines procedures for equipment that interconnect GSTN networks with IP networks to provide satisfactory, transparent delivery of modulated voice-band data (VBD) as encoded audio content over IP (data modems, facsimile terminals and text telephones).

This Recommendation is complementary to the modem relay and voice-band data ITU-T Recs V.150.0 and V.150.1.

Source

ITU-T Recommendation V.152 was approved on 8 January 2005 by ITU-T Study Group 16 (2005-2008) under the ITU-T Recommendation A.8 procedure.

Keywords

Echo canceller, Facsimile over IP, gateway, Internet gateway, Internet Protocol, IP gateway, media gateway, media gateway controller, modem over IP, quality of service, speech coding, TDM, TDM-IP gateway, Text over IP, Textphone over IP, Text Telephone, VBD, voice-band data, voice gateway, Voice over IP,VoIP.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T 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 World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

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

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

CONTENTS

ITU-T Recommendation V.152

Procedures for supporting voice-band data over IP networks

1 Scope

This Recommendation describes the voice-band data (VBD) operation of Voice-over-Internet Protocol (VoIP) gateways and media gateways. The term 'VBD' refers only to the use of suitable voice-band codecs for the transport of data payloads via RTP. The VBD procedures described in this Recommendation shall apply to VBD-only capable gateways. A V.152 gateway only has the guarantee of inter-working with another gateway if that gateway also supports V.152

The negotiation of a VBD-capability does not exclude from a VoIP session any other capabilities such as the transport of audio signals, RFC 2833-based telephone events, ITU-T Rec. T.38 facsimile relay, RFC 2793 text relay and ITU-T Rec. V.150.1 modem relay, etc.

Declaration of VBD support using SDP is detailed in 7.1.

Declaration of VBD support using H.245 is detailed in 7.2.

This Recommendation supports hybrid modes of operation; for example, a device may support a VBD and facsimile relay capability but not a modem relay or text relay capability. In this example, of hybrid operation, modem and text payloads are transported in the VBD mode, while facsimile payloads could be transported in the T.38 facsimile relay mode or in the VBD mode. The negotiation of such hybrid sets of capabilities follows SDP and H.245 mechanisms (clause 11).

This Recommendation describes the default mechanism for transitioning into VBD mode via payload type switching as described in clause 10 and the optional mechanism of utilizing the SSE messages described in clause 11.

2 References

2.1 Normative references

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is published regularly. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

– ITU-T Recommendation G.168 (2004), *Digital network echo cancellers*.

– ITU-T Recommendation G.701 (1993), *Vocabulary of digital transmission and multiplexing, and pulse code modulation (PCM) terms*.

– ITU-T Recommendation G.711 (1988), *Pulse code modulation (PCM) of voice frequencies*.

– ITU-T Recommendation G.726 (1990), *40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)*.

– ITU-T Recommendation G.729 (1996), *Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)*.

– ITU-T Recommendation H.245 (2005), *Control protocol for multimedia communication*.

– ITU-T Recommendation H.248.1v2 (2002), *Gateway control protocol: Version 2*.

– ITU-T Recommendation H.323 (2003), *Packet-based multimedia communications systems*.

– ITU-T Recommendation T.38 (2004), *Procedures for real-time Group 3 facsimile communication over IP networks*.

– ITU-T Recommendation T.120 (1996), *Data protocols for multimedia conferencing*.

– ITU-T Recommendation V.18 (2000), *Operational and interworking requirements for DCEs operating in the text telephone mode*.

– ITU-T Recommendation V.150.1 (2003), *Modem-over-IP networks: Procedures for the end-to-end connection of V-series DCEs*.

– IETF RFC 768 (1980), *User Datagram Protocol*.

– IETF RFC 791 (1981), *Internet Protocol DARPA Internet Program Protocol Specification.*

– IETF RFC 2198 (1997), *RTP Payload for Redundant Audio Data*.

– IETF RFC 2327(1998), *SDP: Session Description Protocol*.

– IETF RFC 2543 (1999), *SIP: Session Initiation Protocol*.

– IETF RFC 2733 (1999), *An RTP Payload Format for Generic Forward Error Correction*.

– IETF RFC 2833 (2000), *RTP Payload for DTMF Digits, Telephony Tones and Telephony Signals*.

– IETF RFC 3264 (2002), *An Offer/Answer Model with the Session Description Protocol (SDP)*.

– IETF RFC 3388 (2002), *Grouping of Media Lines in the Session Description Protocol (SDP)*.

– IETF RFC 3389 (2002), *Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN)*.

– IETF RFC 3550 (2003), *RTP: A Transport Protocol for Real-Time Applications*.

2.2 Informative references and bibliography

– IETF RFC 2234 (1997), *Augmented BNF for Syntax Specifications: ABNF*.

– IETF RFC 3389 (2002), *Real-time Transport Protocol (RTP) Payload for Comfort Noise (CN)*.

– IETF RFC 3711 (2004), *The Secure Real-time Transport Protocol (SRTP)*.

3 Terms and Definitions

This Recommendation defines the following terms.

For terms and definitions not appearing in this clause, see ITU-T Rec. G.701 (1993), *Vocabulary of digital transmission and multiplexing, and pulse code modulation (PCM) terms*.

3.1 general switched telephone network (GSTN): This network includes ATM, PSTN, ISDN, wireless networks and private networks.

3.2 H.248 gateway: A Media Gateway that complies with the ITU-T H.248x series of Recommendations.

3.3 media gateway (MG): The media gateway converts media provided in one type of network to the format required in another type of network. For example, a MG could terminate bearer channels from a switched circuit network (e.g., DS0s) and media streams from a packet network (e.g., RTP streams in an IP network). This gateway may be capable of processing audio, video and T.120 multimedia signals alone or in any combination, and will be capable of full duplex media translations. The MG may also play audio/video messages and performs other interactive voice response (IVR) functions, or may perform media conferencing. For the purpose of this Recommendation, the term media gateway refers to a voice gateway.

3.4 media gateway controller (MGC): Controls the parts of the call state that pertain to connection control for media channels in a media gateway.

3.5 modem relay: The transport of modem data across a packet network using modem termination at the gateways.

3.6 MoIP gateway: A Media Gateway that is compliant with the ITU-T V.150 series of Recommendations.

3.7 VBD gateway: A Media Gateway that is compliant with this Recommendation.

3.8 off-ramp gateway: The IP network access point that calls an Answering DCE. (Abbreviated to G2.)

3.9 on-ramp gateway: The access point that is called by an originating DCE that interfaces to the IP network. (Abbreviated to G1.)

3.10 transcoding: Translation from one type of encoded media format to another different media format (examples: G.711 A-law to μ-law or vice versa, G.711 codec to G.726-40K, G.711 to a broadband codec that operates at 256 kbit/s, etc.).

3.11 audio mode: In this mode, the channel processes speech signals. The mode may include the use of compression algorithms and other processing functions that are not suitable for the transport of modem or facsimile signals.

3.12 voice-band data mode: Is the transport of voice-band data over a voice channel of a packet network with the encoding appropriate for modem signals as defined in clause 6.

3.13 modem: The term modem in this Recommendation covers all V-series modems and Text Telephones types covered in the Annexes of ITU-T Rec. V.18.

4 Abbreviations

This Recommendation uses the following abbreviations.

ANS	Answer Tone (as per ITU-T Rec. V.25)
ASNam	Answer Tone (as per ITU-T Rec. V.8)
/ANS	Answer tone with phase reversals (as per ITU-T Rec. V.25)
ASN.1	Abstract Syntax Notation One
ABNF	Augmented Backus-Naur Form (defined in IETF RFC 2234)
CED	Facsimile CallED tone (defined in ITU-T Rec. T.30)
CI	Call Indicator Signal (per ITU-T Rec. V.8)
CNG	Facsimile Calling tone (per ITU-T Rec. T.30)
DS0	Digital Signal, level 0
DTMF	Dual Tone Multi-Frequency
FAX	Facsimile
FEC	Forward Error Correction
FoIP	Facsimile over Internet Protocol
G3FE	Group 3 Facsimile Equipment
GSTN	General Switched Telephone Network

IETF	Internet Engineering Task Force
IP	Internet Protocol
ITU	International Telecommunication Union
IVR	Interactive Voice Response
MG	Media Gateway
MGC	Media Gateway Controller
MoIP	Modem over Internet Protocol
OLC	Open Logical Channel
PCMU	Pulse Code Modulation μ-law
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
SCN	Switched Circuit Network
SDP	Session Description Protocol
SIP	Session Initiation Protocol
SS7	Signalling System No. 7
TDM	Time Division Multiplex(ing)
ToIP	Text telephony over Internet Protocol
UDP	User Datagram Protocol
VBD	Voice-band Data
VoIP	Voice over Internet Protocol
V21-Preamble	Preamble (defined in 5.3.1/T.30)

5 Conventions

An ITU-T Recommendation, by definition, is not mandatory – compliance is voluntary. The use of the words "shall" and "must" and their negatives "shall not" and "must not" are to be used with care and sparingly. These words are only to be used to express mandatory provisions, when necessary, to give the Recommendation meaning and effect; i.e., if certain values and/or parts of a Recommendation are essential, and the Recommendation will have no meaning if these values and/or parts are not strictly respected or adhered to. Compliance with the Recommendation is achieved only when all mandatory provisions are met. However, the inclusion of mandatory provisions in a Recommendation does not imply, of itself, that compliance with the Recommendation is required of any party.

5.1 Recommendation version

For the purposes of forward and backward compatibility, this Recommendation is assigned a version number, which is defined here.

NOTE – The reader is encouraged to check on the ITU-T website for any normative or informative amendments to this Recommendation.

Version: 1

6 Definition of the VBD mode of operation

Voice-band data is the transport of modem, facsimile, and text telephony signals over a voice channel of a packet network with a codec appropriate for such signals.

For voice-band data (VBD) mode of operation, all voice-band modulated signal samples shall be transported across an IP network using the RTP protocol defined in IETF RFC 3550.

When in VBD mode, a V.152 compliant implementation shall:

- Use a codec that passes voice-band modulated signals with minimal distortion. This codec shall be assigned as the VBD codec with a specific RTP payload type which shall be negotiated with the remote V.152 implementations as described in clause 7.
- Have end-to-end constant latency.
- Disable voice activity detection and comfort noise generation during the data transfer phase.
- Disable any DC removal filters that may be integral with the speech encoder used.

And should consider the appropriate application of:

- The use of echo cancellers on the VBD channel, as per ITU-T Rec. G.168.
- Forward Error Correction (FEC) (e.g., RFC 2733) or other forms of redundancy (e.g., RFC 2198) only if support has been successfully negotiated with the remote V.152 implementation.
- Voice packet loss concealment techniques and algorithms that are suitable for modem and facsimile modulations.

6.1 Minimum requirements for VBD mode of operation

For purposes of interoperability, a V.152 compliant implementation shall support at least both G.711 A-law and G.711 μ-law codecs as VBD codecs.

When negotiating the VBD codec, the initiating V.152 implementation must include in the offer either PCMA or PCMU (or both) in the list of VBD codecs, though other VBD codecs may be additionally specified. The V.152 implementation answering the offer must indicate support for at least one VBD codec, which need not be PCM-based.

Redundancy as per IETF RFC2198 and forward error correction as per IETF RFC2733 are supported options.

7 Negotiation of support of VBD and selection of VBD codec and other VBD enhanced functionality

Negotiation of the support and use of VBD data mode, as defined in this Recommendation, is carried out at call establishment during the initial exchange of the call capabilities of the endpoints establishing the call. Indication of such support entails assigning RTP payload types to VBD as well as the codecs.

The mechanisms for negotiation vary depending on the endpoint's capabilities exchange protocols used, which can be the session description protocol (defined in IETF RFC 2327) or ITU-T Rec. H.245; the call control protocol, such as those defined in ITU-T Rec. H.323, and the session initiation protocol (SIP), defined in IETF RFC 3261; and/or the media gateway control protocols such as defined in ITU-T Recs H.248 and J.171.

This clause shall describe negotiation procedures for mechanisms that use:

- Session description protocol (SDP), defined in IETF RFC 2327), such as, but not limited to, SIP terminals/gateways and H.248 gateways;
- ITU-T Rec. H.245, that complies with ITU-T Rec. H.323.

This Recommendation does not preclude the gateways from negotiating support of other mechanisms such as IETF RFC 2833 telephone-events, ITU-T Rec. T.38, ITU-T Rec. V.150.1 and/or text relay, for transporting non-voice signals. RTP shall be used for the transport of VBD.

7.1 Negotiation using Session Description Protocol (SDP)

For implementations that use the session description protocol the 'gpmd' (general-purpose media descriptor) attribute shall be used to associate payload types in a media information ('m') line with VBD mode. The general form of this attribute line is:

```
a=gpmd:<format> <parameter list>
```

In the context of VBD declaration, the <format> must be an RTP/AVP payload type. The <parameter list> is a semicolon-separated list of "parameter=value" pairs. For RTP/AVP formats, these pairs address parameters that are not part of their standard MIME definition. For sessions supporting this Recommendation, the parameter of interest is the Boolean 'vbd' that may have the value of 'yes' or 'no'. When set to 'yes' the attribute indicates that the implementation supports VBD mode as described in this Recommendation.

Omission of the 'gpmd' attribute with a "vbd=yes" attribute/value pair for any codec in the SDP session description shall be construed as non-support of VBD mode operation as defined in this Recommendation.

The payload type marked for voice-band data (VBD) treatment should be a dynamic payload type. It is possible that a codec, such as PCMU, be declared with both static and dynamic payload types, with only one of the two marked for voice-band data use (see example 1 below). If a codec, such a PCMU or PCMA, is declared with only a static payload type, and is also marked for voice-band data use, then this codec must not be used for carrying voice (see example 2 below).

In addition to negotiating support of V.152 and the corresponding RTP payload type, a V.152 implementation should include the 'maxmptime' attribute (maximum multiple ptime) to indicate the supported packetization period for all codec payload types.

```
a=maxmptime:<list of packet times separated by space>
```

This attribute is a media-level attribute. The maxmptime attribute defines a list of maximum packetization time values, expressed in milliseconds, the endpoint is capable of using (sending and receiving) for this connection. There shall be precisely one entry in the list for each <format> entry provided in the "m=" line. Each entry is separated by a space. Entry number j in this list defines the maximum packetization time for entry number j in the "m=" line. The first entry in the list shall be a decimal number whereas subsequent entries in the list shall be either a decimal number or a hyphen. For those media formats where a single maximum packetization rate does not apply (e.g., non-voice codecs such as telephone-event or comfort noise), a hyphen ("-") shall be included at the corresponding location in the list of packetization periods.

When receiving an SDP session description, the maxmptime attribute conveys the list of maximum packetization periods that the remote endpoint is capable of using for this connection; one for each media format in the "m=" line. For media formats whose packetization period is specified as a hyphen ("-"), the VBD gateway shall use one of the maximum packetization periods that was actually specified in the list.

The "a=ptime" attribute, defined in RFC 2327, shall be ignored if the SDP session description contains the "maxmptime" attribute.

If the "maxmptime" attribute is absent, then the value of the "ptime" attribute, if present, shall be taken as indicating the packetization period for all codecs present in the "m=" line.

If neither the 'ptime' nor 'maxmptime' attribute are present in the SDP session description, then a V.152 implementation shall assume the default packetization period defined in RFC 3550 (which is 20 ms for G.711 and G.726-32k). A V.152 implementation shall not transmit V.152 packets with a packetizaton time greater than the one offered by the remote end.

Below is an SDP-related example that indicates support of V.152 as per this Recommendation. For clarity purposes, the example only shows the media descriptions of the SDP session description.

Example 1

```
m=audio 3456 RTP/AVP 18 0 13 96 98 99
a=maxmptime:10 10 - - 20 20
a=rtpmap:96 telephone-event/8000
a=fmtp:96 0-15, 34, 35
a=rtpmap:98 PCMU/8000
a=gpmd:98 vbd=yes
a=rtpmap:99 G726-32/8000
a=gpmd:99 vbd=yes
```

In the example directly above, static payload type '0' and dynamic payload type '98' each represent the encoding format 'PCMU'. The payload type '0' is not associated with VBD. The payload types '98' (PCMU) and '99' (32 kbit/s ADPCM) are, however, associated with VBD. Concerning the maximum packetization times for each payload type: Voice packets use 10 ms, VBD packets use 20, and a dash is assigned to payload types 13 (silence indication packets) and 96 (RFC 2833 packets) indicating that a maximum ptime is not applicable or necessary.

Example 2

```
m=audio 3456 RTP/AVP 0 18 98
a=gpmd:0 vbd=yes
a=rtpmap:98 G726-32/8000
a=gpmd:98 vbd=yes
a=ptime:20
```

In this example, the static payload type of 0 (PCMU) is marked for VBD treatment, along with the dynamic payload type '98' (mapped to 32 kbit/s ADPCM). Thus, payload type 0 must not be used for carrying voice. It also indicates that the VBD gateway can receive voice and VBD packets with a size of 20 ms.

NOTE – The use of static payload types for VBD is strongly discouraged because there is a danger that a non-V.152 system would see a proposal with, say, G.711 VBD and G.729. However, not understanding the VBD attributes, it may consider G.711 as a valid audio codec. However, network operators may prefer that G.711 not be utilized, except in the case where VBD is necessary, and that all voice shall be G.729. To illustrate this point, consider this offer:

```
m=audio 15400 RTP/AVP 0 18
a=gpmd:0 vbd=yes
```

and consider this answer:

```
m=audio 15400 RTP/AVP 0 18
```

The systems would then, most likely, communicate using G.711, rather than the intended G.729, for voice.

7.1.1 Mechanism for indicating support of V.152 using H.248/Megaco

Under H.248, the Media Gateway Controller (MGC) uses Local and Remote descriptors to reserve and commit MG resources for media decoding and encoding for the given Stream(s) and Termination to which they apply. The MG includes these descriptors in its response to indicate what it is actually prepared to support. When text encoding the protocol, the descriptors consist of SDP session descriptions that describe the call capabilities.

Support of V.152 shall only be applied on Ephemeral terminations, via the Local and/or Remote descriptors.

For an MG to reserve and commit resources for more than one call capability alternative, the MGC must set the ReserveGroup and ReserveValue properties of the LocalControlDescriptor to 'True'.

Thus, if a list of payload types is offered in a Local and/or Remote descriptor, such as the following example 3 of an Add ephemeral termination command illustrates (note similar applies if the command was a Modify or Move), the media gateway will select from the list only those payloads for which it can reserve and commit resources and shall send a reply to the MGC containing the alternatives for the Local and/or Remote descriptor that it selected, as described in ITU-T Rec. H.248.1:

Example 3a

```
MGC to MG:
MEGACO/1.0 [123.123.123.4]:55555
Transaction = 11 {
  Context = $ {
        Add = $ {
      Media {
        Stream = 1 {
          LocalControl { Mode = ReceiveOnly, ReserveGroup = True,
ReserveValue = True},
          Local {
v=0
c=IN IP4 $
m=audio $ RTP/AVP 18 0 98 99
a=rtpmap:98 PCMU/8000
a=gpmd:98 vbd=yes
a=rtpmap:99 G726-32/8000
a=gpmd:99 vbd=yes

           }; IP termination for audio and VBD
         }
       }
     }
  }
}
```

Alternatively, an MGC may leave it up to the MG whether it wants to indicate that it supports VBD, as per this Recommendation, and to select its dynamic payload type for VBD mode of operation by including CHOOSE (i.e., $) in the payload type list field as example 3a illustrates:

Example 3b

```
MGC to MG:
MEGACO/1.0 [123.123.123.4]:55555
Transaction = 11 {
  Context = $ {
        Add = $ {
      Media {
        Stream = 1 {
          LocalControl { Mode = ReceiveOnly, ReserveGroup = True,
ReserveValue = True},
          Local {
v=0
c=IN IP4 $
```

```
m=audio $ RTP/AVP 18 0 $
           }; IP termination for audio and VBD
         }
       }
     }
  }
}

MG to MGC response:
MEGACO/1.0 [123.123.123.4]:55555
Transaction = 11 {
  Context = 34444 {
        Add = Te/1 {
      Media {
        Stream = 1 {
          LocalControl { Mode = ReceiveOnly, ReserveGroup = True,
ReserveValue = True},
          Local {
v=0
c=IN IP4 $
m=audio $ RTP/AVP 18 0 98 99
a=rtpmap:98 PCMU/8000
a=gpmd:98 vbd=yes
a=rtpmap:99 G726-32/8000
a=gpmd:99 vbd=yes
           }; IP termination for audio and VBD
         }
       }
     }
  }
}
```

Once an MG has acknowledged a set of call capability alternatives, the MG is requested to reserve resources so that it can decode or encode the media stream according to any of the alternatives. Thus, in the above example 3a, if the MG supports G.729 and G.711 for audio and G.711 for VBD, (as per this Recommendation), then, in accordance with ITU-T Rec. H.248.1, the MG must reserve resources such that it can decode one RTP stream in any of the formats in its response ant any time during the call, i.e., G.711 audio format, G.729 audio format or G.711 VBD format.

If a specific relay mechanism (e.g., T.38, V.150.1, etc.) is indicated as the preferred mechanism above that of VBD, then, for the applicable devices, the relay mechanisms shall be used instead of VBD. For example, if a remote descriptor indicates T.38 as preferred over VBD, then an MG shall use T.38 for all G3FE equipment instead of VBD.

If an MG cannot guarantee that it can commit and reserve the resources for VBD for the call being set up, then in accordance with ITU-T Rec. H.248.1 it shall not include the 'gpmd' attribute (that indicates support of V.152) in its response SDP session description.

Note that this mechanism does not preclude a H.248 MG implementation from sending to the MGC observedEvents indicating signals detected, as described in the H.248.2 package.

7.1.2 Mechanism for indicating support of V.152 using SIP

In the case of SIP terminals, the offer/answer model described in IETF RFC 3264 shall be used to earmark one or more RTP payload types for VBD operation as defined in this Recommendation.

Just as a SIP-compliant terminal would indicate support of more than one audio codec payload or support of other payload types (e.g., RFC 2833 for DTMF relay) within a media stream, a SIP-compliant implementation shall indicate support of V.152 by including the payload types as described in 7.1.

If multiple media-descriptions are being offered and if the implementations cannot support simultaneous reception and transmission of the various media types, then the 'group', 'mid' and 'FID' attributes described in IETF RFC 3388 shall be used to indicate alternative support of each of the offered media types (as illustrated below in example 5).

Once a gateway has indicated support of V.152 in addition to other mechanisms within an SDP session description (such as but not limited to audio, Facsimile relay via ITU-T Rec. T.38 DTMF relay as per IETF RFC 2833, etc.), the gateway shall be capable of switching between any of the supported, and mutually negotiated, RTP payload types, at any time during a call.

7.1.2.1 Mechanism for indicating preference of VoIP relay mechanisms above VBD

SIP currently does not have a mechanism for indicating in a clear manner that a gateway would like to use a specific relay mechanism (e.g., T.38, V.150.1, Text Relay) instead of VBD. Hence, this clause defines the syntax and use of an attribute that indicates a preference list of modem and facsimile transport methods in a V.152 implementation that supports any of the following alternative transport methods:

- Facsimile Relay of IP via ITU-T Rec. T.38;
- Modem Relay over IP via ITU-T Rec. V.150.1;
- Text relay.

The attribute is called the 'pmft' attribute and its formatting in the SDP session description is described by the following ABNF syntax:

```
pmft-attribute          =   "a=pmft:" *(SPACE modem-fax-transport)
modem-fax-transport         =   1* ("V1501" / "T38" / "V151")
```

This attribute allows a V.152 implementation to indicate whether it prefers any of the listed relay transport mechanism above VBD mode. Omission of this attribute in an SDP session description means that VBD mode is the preferred transport mechanism of voice-band data.

When included in an SDP session description, this attribute shall always be placed at the session level.

For example, a V.152 implementation that also supports V.150.1 for modems and T.38 for Facsimile, and prefers to use these relay mechanisms whenever possible instead of VBD, shall include, at the session level of the SDP session description, the following 'pmft' attribute:

```
a=pmft: T38 V1501
```

A V.152 implementation that receives the above 'pmft' attribute, and is able to support both the relay mechanisms specified in this example, shall include the same 'pmft' attribute in its response. Thus, when the call is set up, all G3FE shall be transported via T.38, voice-band modems that are supported by V.150.1 shall be transported via V.150.1 and all other modems (e.g., text telephones) shall be transported by V.152.

A gateway answering an SDP session description offer that includes the 'pmft' attribute, if indicating preference of the supported relay mechanism over VBD, shall include in the response SDP session description the 'pmft' attribute with the relay mechanism specified. If a relay mechanism is not supported, then that relay mechanism shall be removed from the list of the pmft attribute.

Once a specific relay mechanism (e.g., T.38, V.150.1, etc) is indicated as the preferred mechanism above that of VBD, such relay mechanisms shall be used instead of VBD.

A gateway answering an SDP session description offer that shows the capability to perform relay mechanisms but does not include the 'pmft' attribute may include in the response SDP session description the 'pmft' attribute with the relay mechanisms specified to show a preference to use these. For example: if the initial SDP session description offer from a gateway that only supports V.152 and T.38 does not include the 'pmft' attribute because it prefers to use VBD above T.38, then it would include an SDP session description such as:

```
v=0
o=Offerer  0 0 IN IPV4 <IPAdressA>
s=-
t=0 0
p=+1
c=IN IP4  <IPAddressA>
a=group:FID 1 2
m=audio <udpPort x> RTP/AVP 18 0 13 96
a=mid:1
a=ptime:10
a=rtpmap:96 PCMU/8000
a=gpmd: 96 vbd=yes
m=image <udpPort y> udptl t38
a=mid:2
a=T38version:0
a=T38FaxRateManagement:transferredTCF
a=T38FaxUdpEC:t38UDPRedundancy
```

However, the answerer gateway (that may not be in as reliable a network) may always prefer T.38 above that of VBD for facsimile transmission. Thus, it would include in its response the following:

```
v=0
o=Answerer  0 0 IN IPV4 <IPAdressB>
s=-
t=0 0
p=+1
c=IN IP4  <IPAddressB>
a=group:FID 1 2
a=pmft: T38
m=audio <udpPort x> RTP/AVP 18 0 13 96
a=mid:1
a=ptime:10
a=rtpmap:96 PCMU/8000
a=gpmd: 96 vbd=yes
m=image <udpPort y> udptl t38
a=mid:2
a=T38version:0
a=T38FaxRateManagement:transferredTCF
a=T38FaxUdpEC:t38UDPRedundancy
```

The offerer gateway on receipt of such answer shall transport G3FE data using T.38 and all other modems, or non-G3FE must be transported via V.152.

7.1.3 Examples of indicating support for V.152 using the session description protocol

This clause shall provide a few examples of SDP session descriptions sent by implementations that support V.152 in addition to other Recommendations (such as, but not limited to, Voice, T.38, ToIP, and V.150.1).

Example 4: An implementation that supports V.152 (using the dynamic payload type 96 and G.711 mμ-law as the VBD codec) and the voice codecs G.711 mμ-law, silence suppression and G.729, shall transmit the following SDP session description, only those lines that are relevant to this Recommendation are highlighted in bold:

```
v=0
o=-  0 0 IN IPV4 <IPAdressA>
s=-
t=0 0
p=+1
```

```
c=IN IP4  <IPAddressA
m=audio <udpPort A> RTP/AVP 18 0 13 96
a=ptime:10
a=rtpmap:96 PCMU/8000
a=gpmd: 96 vbd=yes
```

A V.152 implementation that receives an SDP session description, as in the above example, shall interpret it as the remote gateway's capability to support V.152 and that the payload type that shall be used for VBD packets is 96.

Example 5: A call is set up between Gateway A that supports V.152, T.38, IETF RFC 3389 silence suppression and Voice codec G.729 and PCMU, and a Gateway B that supports T.38, silence suppression and Voice codecs G.729 and PCMU but does not support V.152.

The SDP transmitted by gateway A will be of the form:

```
v=0
o=GatewayA  0 0 IN IPV4 <IPAdressA>
s=-
t=0 0
p=+1
c=IN IP4  <IPAddressA>
a=group:FID 1 2
m=audio <udpPort x> RTP/AVP 18 0 13 96
a=mid:1
a=ptime:10
a=rtpmap:96 PCMU/8000
a=gpmd: 96 vbd=yes
m=image <udpPort y> udptl t38
a=mid:2
a=T38version:0
a=T38FaxRateManagement:transferredTCF
a=T38FaxUdpEC:t38UDPRedundancy

(…….additional T.38 attributes may follow…..)
```

Gateway B, which does not support VBD, shall respond with an SDP that omits all reference to V.152:

```
v=0
o=GatewayB  0 0 IN IPV4 <IPAdressB>
s=-
t=0 0
p=+1
c=IN IP4  <IPAddressB>
a=group:FID 1 2
m=audio <udpPort w> RTP/AVP 18 0 13
a=mid:1
a=ptime:10
m=image <udpPort z> udptl t38
a=mid:2
a=T38version:0
a=T38FaxRateManagement:transferredTCF
a=T38FaxUdpEC:t38UDPRedundancy
(…….additional T.38 attributes may follow…..)
```

On receipt of the above SDP, gateway A shall understand that gateway B does not perform V.152. Thus gateway A shall not transition into a VBD mode.

Example 6: Gateway A supports Voice Codec G.729, V.152 and V.150.1. Gateway B also supports V.150.1 and Voice Codec G.729 but does not support V.152:

NOTE – This example 6 shows the minimum number of lines needed to construct an SDP-compliant session descriptor that includes all attributes that are mandatory for the representation of SPRT modem relay and V.152.

```
SDP from Gateway A:
     v=0
     o=Gateway A 25678 753849 IN IP4 128.96.41.1
     s=
     c=IN IP4 128.96.41.1
     t=0 0
     m=audio 49230 RTP/AVP 0 8 18 97 98
     a=gpmd:0 vbd=yes
     a=gpmd:8 vbd=yes
     a=rtpmap:97 telephone-event/8000
     a=fmtp:97 0-15,32,33,34,35,66,70
     a=rtpmap:98 v150fw/8000
     m=audio 49232 udpsprt 100
     a=sprtmap:100 v150mr/8000
     a=fmtp:100 mr=0; mg=1;DSCselect=3;mrmods=1,2;jmdelay=no;versn=1.1
```

In this example, ports 49230 and 49232 are used for the RTP/AVP and SPRT media streams respectively. Within the RTP/AVP media stream, the static payload types of 0 (PCMU) and 8 (PCMA) are marked for VBD treatment via the 'gpmd' attribute, and thus cannot be used for Voice.

Also note, in accordance with SIP, the above SDP implies simultaneous support of audio 'rtp/avp' and audio 'udpsprt'. To indicate that only one media type can be supported at a time, the 'group' attribute with the FID semantics, together with the 'mid' attribute, should be used as specified in RFC 3388 (see example 5).

Gateway B does not support V.150.1 (implementations which also must support VBD), Gateway B shall respond with an SDP as follows:

```
     v=0
     o=GatewayB 25678 753849 IN IP4 128.96.41.1
     s=
     c=IN IP4 128.96.41.1
     t=0 0
     m=audio 49230 RTP/AVP 0 8 18 97 98
     a=gpmd:0 vbd=yes
     a=gpmd:8 vbd=yes
     a=rtpmap:97 telephone-event/8000
     a=fmtp:97 0-15,32,33,34,35,66,70
     a=rtpmap:98 v150fw/8000
     m=audio 49232 udpsprt 100
     a=sprtmap:100 v150mr/8000
     a=fmtp:100 mr=0; mg=1;DSCselect=3;mrmods=1,2;jmdelay=no;versn=1.1
```

Because both gateways have negotiated support of SSE, they shall use SSEs to indicate transition between voice and VBD.

7.1.4 Optional V.152 capabilities

This clause describes the SDP representation of information that may be optionally declared at session establishment time. The absence of their declaration shall be construed by a V.152 implementation as an indication that the remote V.152 implementation does not support them.

7.1.4.1 Declaration of redundancy and forward error correction

The declaration, in SDP, of RFC 2198 redundancy and RFC 2733 FEC, shall conform to the rules in the applicable IETF source documents. When supporting text telephones, on networks where the error character service requirements as specified in Annex A.3/F.700 is exceeded due to packet loss, then this Recommendation strongly encourages the appropriate use of IETF RFC 2198 redundancy and IETF RFC 2733 FEC for the IP network to which it is attached. However, on some networks the application of redundancy/FEC may contribute to the character error rate and should not be used.

Although the RFC 2198 rules are not repeated here, declaring RFC 2198 support with a redundancy level of 3 for a VBD codec is illustrated with an example:

```
m=audio 3456 RTP/AVP 0 15 102
a=gpmd:0 vbd=yes
a=rtpmap:102 red/8000
a=fmtp:102 0/0/0/0
```

Examples of the declaration of FEC support are found in RFC 2733. This includes the use of a separate FEC stream and the combination of the FEC stream with the primary stream via RFC 2198 encapsulation. In the case when FEC is a separate stream, RFC 2733 uses an 'fmtp' line to associate this stream with an IP address and port. When FEC packets are sent to the same IP address and port (albeit a different SSRC) as the media packets they qualify, there is no need for the 'fmtp' line to associate the 'parityfec' payload type with an IP address and port. Thus, in the following SDP segment:

```
c=IN IP4 224.2.17.12
t=0 0
m=audio 49170 RTP/AVP 0 15 78
a=gpmd:0 vbd=yes
a=rtpmap:78 parityfec/8000
a=fmtp:78 49170 IN IP4 224.2.17.12
```

The last line is superfluous and may be omitted. Likewise, the absence of an 'fmtp' line associating an IP address and port with a FEC payload type WILL be construed to mean that the FEC packets are to be sent to the same IP address and port as the media packets they qualify.

7.1.4.2 Optional vendor – Specific parameters

The 'vndpar' (vendor parameters) attribute may be used to declare vendor codes for coordinating enhanced operation over and above those indicated in ITU-T Rec. V.152. It shall be possible to safely ignore vendor-specific parameters and still maintain interoperability with equipment conforming to this Recommendation. Hence, proprietary enhancements cannot be a substitute for the basic features required for compliance with this Recommendation.

The format of the 'vndpar' attribute line is as follows:

```
a=vndpar:<vendorIDformat> <vendorID> <vendorSpecificDataTag>
              [<vendorSpecificData>]
```

The <vendorIDformat>, a decimal, indicates the format of the following <vendorID> field. The following values are defined:

Integer representation	Vendor ID format
1	ITU-T Rec. T.35
2	IANA Private enterprise number

The <vendorID> may be represented in hex or in decimal format. If represented in hex, it is has a '0x' prefix. Generally, if the vendor ID format is ITU-T T.35, the hexadecimal format is preferred. If it is the IANA private enterprise number (http://www.iana.org/assignments/enterprise-numbers), the decimal format is preferred.

When the vendor ID format is T.35, the vendor ID consists of a country code followed by a vendor code. The country code consists of four octets and the vendor ID consists of two octets. If the representation of the vendor ID is hexadecimal, leading zeros in the country code may be omitted, while leading zeros in the vendor code may not be omitted.

When the <vendorID> is the vendor's private enterprise number, leading zeros may be omitted.

The <vendorSpecificDataTag> is a decimal integer between 0-255. If used, values in the range 1-255 are uniquely mapped, via the 'vndpar' attribute, to the combination of the vendor specified in the <vendorID> and the proprietary capabilities indicated by <vendorSpecificData>. This mapping, which exists for the duration of a session, does not persist across sessions. Further, each side may choose this integer independently of the other end. Due to the compactness of this index, a gateway or endpoint may use it in a number of places. A value of 0 is a null value. When present, it is equivalent to omitting the <vendorSpecificDataTag>. A null value of the <vendorSpecificDataTag> is not associated with any vendor ID.

It shall be possible for an endpoint or gateway to declare multiple (1-255) 'vndpar' attribute lines in an SDP session description. Each of these lines may indicate a different vendor. In addition, multiple 'vndpar' lines may indicate the same vendor. When multiple 'vndpar' lines are declared in an SDP session descriptor, each value of <vendorSpecificDataTag> must either be unique within all 'vndpar' lines in the session descriptor or null (0). If non-null, the <vendorSpecificDataTag> may serve as a dynamically assigned feature identifier for the vendor.

Inclusion of the parameter `<vendorSpecificData>` is optional. When included, this is a vendor-defined octet string consisting of one or more octets. Since it consists of an integer number of octets, it is represented by an even number of hex characters. No '0x' prefix is needed. No size limitation is specified since SDP parsers can ignore another vendor's string without checking its length. A vendor is permitted to add additional structure to the `<vendorSpecificData>` field such that features are identified by their position in this field. A vendor may also elect to add explicit feature identification within the `<vendorSpecificData>` field. When present, these supplement the <vendorSpecificDataTag>.

Note that the vendor is not precluded from using the `<vendorSpecificData>` field to communicate parameters that are not related to V.152.

7.2 Use of VBD in H.323 systems

H.323 systems support V.152 through the use of the **VBDCapability** capability defined in ITU-T Rec. H.245. This capability, which is a type of **AudioCapability**, is used during capability exchange and in the open logical channel (OLC) signalling to indicate support for VBD channels and to signal the opening of those channels. Since VBD media flows are generally switched within a single RTP session with normal voice audio and other audio-related media (e.g., RFC 2833), OLC proposals and terminal capability set messages generally utilize the multiple payload stream (MPS) constructs in ITU-T Rec. H.245.

7.2.1 Fast connect procedures

H.323 systems may offer one or more logical channel proposals in the SETUP message transmitted to the called party. The H.323 device orders those logical channel proposals in order of preference. This allows an endpoint to indicate its preferred mode of operation and allows the called device to understand what is preferred, but to also accept any alternative modes offered by the calling device.

If the calling device prefers to use VBD for the transport of all voice-band data, including Facsimile, text, and modem signalling, the first proposal in the OLC would consist of a non-VBD audio codec and a VBD codec. If the calling endpoint also supports T.38 relay over RTP, for example, it might offer as a second proposal a non-VBD audio codec, a VBD codec, and T.38. In this way, the called device knows that the calling device prefers to use VBD for all voice-band data, but is also willing to do T.38 relay over RTP in the case that the called device has this preference. As with normal Fast Connect procedures, the called device is at liberty to accept any of the alternative proposals or refuse all of them and utilize normal H.245 signalling to open logical channels.

Generally, H.323 devices would also signal as additional OLC proposals, ones that contain different audio codecs in combination with VBD codecs. Further, devices would also signal alternatives that offered only a non-VBD codec as the choice for media, in the case that the called device does not support this Recommendation. The choice of OLC proposals, the order of proposals, and the selection is a matter of implementation.

H.323 devices compliant with this Recommendation may also utilize Extended Fast Connect, which allows devices to renegotiate media streams and to make counter-proposals to those OLCs offered by the remote endpoint. Refer to ITU-T Rec. H.460.6 for the procedures related to Extended Fast Connect.

By no means does this Recommendation override the rules defined in ITU-T Rec. H.323 related to Fast Connect procedures or the rules defined in ITU-T Rec. H.460.6 for Extended Fast Connect.

7.2.2 Exchanging VBD capabilities

Devices specify support for VBD by including capabilities of the type **VBDCapability** in the H.245 **TerminalCapabilitySet** message. As with other types of media, these capabilities may be grouped into capability descriptors to signal sets of simultaneous capabilities. Further, since VBD is generally one type of audio that is switched within the same RTP session as other media, **VBDCapability** capabilities are generally only defined as part of a Multiple Payload Stream (MPS). However, since a device may wish to open a VBD stream that does nothing more than transmit media as VBD, capabilities may be defined and used outside of an MPS.

7.2.3 H.245 logical channel signalling procedures

Once H.323 devices have exchanged capabilities, they may open logical channels by transmitting Open Logical Channel (OLC) messages. The procedures for logical channel signalling are defined in ITU-T Rec. H.323 and no additional procedures are defined in this Recommendation.

Since H.323 devices operate asynchronously, it is possible that one device may transmit an OLC message offering one set of capabilities, while the peer device transmits an OLC with an incompatible set of capabilities. For example, one device may propose an OLC that suggests the use of {G.729, VBD/G.711, T.38} while the peer device sends an OLC which suggests the use of {G.723.1, VBD/G.726}. Of course, both messages should be independently legal based on the exchanged capabilities. While H.323 permits devices to use different audio codecs in each direction, it may not be preferred. In this case, the fact that one side proposes to use T.38 over RTP and the other side does not is a problem. In all such cases, H.323 specifies that the master shall resolve such conflicts by rejecting the OLC with a reason of **masterSlaveConflict** or other appropriate reason. Devices should not fail as a result, but should converge to a common mode.

Devices should utilize the request mode message in H.245 in order to suggest a compatible mode of operation. Either the master or the slave device may transmit a request mode message. Note, however, that the request may be rejected. Ultimately, the slave device may have no choice, except to use the preferred mode of the master. Even so, the master should honor requests from the slave device when possible.

As an example to illustrate the opening of a media channel in accordance with this Recommendation, consider an OLC that has a G.729 voice stream, a G.711 A-law VBD stream that is protected with redundancy encoding, an RFC 2833 stream, and a T.38 over RTP stream. The **OpenLogicalChannel** would essentially have a composition similar to what is shown here:

```
{
     forwardLogicalChannelNumber 1,
     forwardLogicalChannelParameters {
          dataType : multiplePayloadStream {
               element {
                    dataType : audioData : g729 2
               },
               element {
                    dataType : redundancyEncoding {
                         primary {
                              dataType : audioData : vbd : g711Alaw64k 160
                         },
                         secondary {
                              {
                                                  dataType : audioData : vbd
                           : g711Alaw64k 160,
                              }
                         }
                    },
                    payloadType 101     -- The PT for the RFC 2198 packet
               },
               element {
                    dataType : audioData : audioTelephonyEvent {
                              audioTelephoneEvent : "0-15,32,33"
                    },
                    payloadType 102
               },
               element {
                    dataType : audioData : genericDataCapability {
                              capabilityIdentifier : standard {
                                   itu-t(0) recommendation(0) t(20) 38
                                   h245-audio-capability(0)
                              },
                         nonCollapsing {
                              {
                                   parameterIdentifier : standard : 0,
                                   parameterValue : booleanArray : 0
                              },
                              {
                                   parameterIdentifier : standard : 1,
                                   parameterValue : unsignedMin : 0
                              },
                              {
                                   parameterIdentifier : standard : 2,
                                   parameterValue : genericParameter
                                   {
                              {
                                   parameterIdentifier : standard : 1,
                                   parameterValue : logical
                              }
                                   }
                              },
                              {
                                   parameterIdentifier : standard : 3,
                                   parameterValue : unsigned32Max : 200
                              },
                              {
                                   parameterIdentifier : standard : 4,
                                   parameterValue : unsigned32Max : 72
                              },
                              }
                         }
                    },
```

```
                    payloadType 103
                }
            }
        },
    multiplexParameters : h2250LogicalChannelParameters {
            sessionID 1
            }
    }
```

8 The use of RFC 2833 modem/facsimile/text telephone events

The declaration of IETF RFC 2833 telephone-events ANS (32), /ANS (33), ANSam (34) and /ANSam (35) is optional. If these events are declared by a media gateway, the remote media gateway may use RFC 2833 to transmit these events. in place of VBD packet transmission. If both media gateways indicate support of the RFC 2833 telephone-events ANS (32), /ANS (33), ANSam (34) and /ANSam (35), then these events shall be used by the media gateways for echo canceller control per ITU-T Rec. G.168. If either end does not indicate this support, then the media gateways shall detect the 2100 Hz tone with phase reversals signal for echo canceller disabling on their incoming VBD packet stream.

When using IETF RFC 2833 telephone-events, the amount of in-band signal leakage into the IP Network for ANS, ANSam, /ANS, and /ANSam signals shall be less than 50 ms.

9 VBD stimuli

This clause lists the stimuli that should be detected, per type of application, by a VBD gateway to initiate a transition to the VBD mode of operation, as described in clause 10.

The list of stimuli below is not exhaustive and there may be other tones that can be used to initiate a transition to VBD for the listed applications:

- *For Facsimile applications*
  - CED as per ITU-T Rec. T.30;
  - ANSam as per ITU-T Rec. V.8;
  - Preamble as per 5.3.1/T.30;
  - CNG as per ITU-T Rec. T.30.
- *For Modem applications*
  - ANS as per ITU-T Rec. V.8;
  - ANSam as per ITU-T Rec. V.8;
  - 2225 Hz answer tone as per Appendix VI/V.150.1;
  - Unscrambled binary ones signal as per ITU-T Rec. V.22;
  - CI signals that precede ANSam, as per ITU-T Rec. V.8;
  - Initiating Segment 1 dual tones (1375 Hz and 2002 Hz) as per ITU-T Rec. V.8 *bis*.
- *For Text Telephony applications*
  - ANS as per ITU-T Rec. V.8;
  - ANSam as per ITU-T Rec. V.8;
  - Text telephone signals as defined by 5.1.1/V.18;
  - DTMF signals only if RFC 2833 telephone-events are not supported;
  - CI signals that precede ANSam, as per ITU-T Rec. V.8;

– Calling Tone (CT) signals that precede ANS, as per ITU-T Rec. V.25;
– Initiating Segment 1 dual tones (1375 Hz and 2002 Hz) as per ITU-T Rec. V.8 *bis*.

In addition to the above list, if any other unrecognized tonal non-voice signal is detected this may be used to transition into VBD mode.

VBD gateways should keep signal leakage to a minimum to prevent erroneous behaviour of end terminals.

10 Procedures for transitioning between audio mode and VBD mode

This clause describes the transitioning mechanism for an implementation that only supports VBD as per this Recommendation and Voice, but does not support any relay mechanisms such as RFC 2833, T.38 or V.150.1, nor VBD as per V.150.1.

The mechanism described in this clause shall be the default mandatory mechanism used by V.152-compliant gateways if no other mechanisms have been successfully negotiated between the gateways, otherwise the mutually negotiated mechanism (such as those described in clause 11) shall be used in preference to this method.

The transition from audio mode to VBD mode is performed when the VBD detectors classify an input signal as VBD.

Detection of the stimuli described in clause 9 shall be carried out at least in the direction from the GSTN to the IP network; however, detection in the direction from the IP network to the GSTN network is not precluded.

On detection of any of the stimuli described in clause 9, if the corresponding IETF RFC 2833 telephone-event has not been mutually negotiated, a V.152 implementation must transmit them as in-band as VBD packets.

If the V.8 CI and the V.8 *bis* signals are transmitted in-band rather than as RFC 2833 events, a shift to VBD must not lose any part of the signals. The choice of using in-band indication or RFC 2833 for indicating these signals is dependant upon capability declaration, whether a VBD channel is available, and the preference of the transmitter.

When in the VBD media state, a media gateway may use RFC 2833 in lieu of voice-band transmission to communicate to the remote gateway any of the voice-band data stimuli indicated in clause 9.

The use of IETF RFC 2833 in this case is contingent on the capabilities declared by the remote gateway.

When in the VBD media state, the unscrambled binary ones signal is communicated in-band. There is no IETF RFC 2833 support for this signal.

When in the VBD media state, the text telephone signals are communicated in-band. The media gateway shall not lose any characters at the onset of in-band VBD transmission.

The gateway shall suppress a voice-band data stimulus from the bearer path if it intends to convey the stimulus as an IETF RFC 2833 telephone event. This shall be done immediately on detection of the stimulus. A media gateway knows, prior to the detection of a voice-band data stimulus, whether it will transmit the stimulus in-band or via an RFC 2833 telephone event. This knowledge is based on the remote gateway's capabilities (whether it can receive an RFC 2833 encoding of that stimulus) and the local gateway's own choice (since it may use in-band transmission regardless of the remote gateway's capability declaration).

Once VBD has been mutually negotiated by the two gateways, using the procedures described in clause 7, a gateway that complies with this Recommendation shall be able to receive and appropriately decode, from the IP network, RTP packets with any of the supported negotiated payload types for a particular call: and, hence, a V.152 implementation shall transition from Voice to VBD on receipt of an RTP packet that has the negotiated VBD payload type.

Additionally, gateways may optimize operation by doing one of the following:

- Loading both the audio and VBD codecs to streamline quick, on-the-fly transitions between talkspurts and textspurts.
- Staying in the VBD mode across talkspurts and textspurts.

Thus, on detection, in the direction from the GSTN to IP network of the appropriate VBD signals, a VBD Gateway shall transition to VBD and transmit as soon as possible RTP packets with the corresponding negotiated VBD payload type. Reception of an RTP packet that has the pre-negotiated VBD payload type at the remote end shall cause a VBD Gateway to transition to VBD mode, but only if prior to receiving the VBD RTP packet it received RTP packets that correspond to the state it was previously in (e.g., Voice packets). The reason for the latter rule is explained in the following example:

Consider two VBD gateways termed A and B connected via an IP network and each having a GSTN network on their other sides. There will be a period during a call when both VBD Gateways are in VBD mode. Gateway A transitions into Audio mode due to detection of voice signals in the direction from the GSTN to the IP network, which will cause it to transition into Voice and transmit Voice RTP packets. While the first transmitted Voice RTP packet is traversing the IP network, the remote end (gateway B) is still transmitting VBD RTP packets, because it has not detected anything on its GSTN side, nor has it yet received the Voice RTP packets. To avoid gateway A from transitioning erroneously back into VBD mode, it must not transition back to VBD until it has first received (the pre-negotiated) Voice RTP packets, which it should expect to receive due to it (i.e., gateway A) transitioning into voice.

NOTE – An implementation shall be able to handle out of order RTP packets (e.g., a Voice packet followed by a VBD packet that was actually sent before the Voice packet).

Transition from VBD to Voice may be carried out by detection:

- In the direction from the GSTN to IP network of any of the following stimuli:
  - end of modem or facsimile signals;
  - voice signals;
  - detection in both directions, GSTN to IP and IP to GSTN, of silence. With the following caveats:
    - For text telephones, the appropriate detection of silence must be considered because text telephone conversations may have long periods of silence.
    - For the case of facsimile calls, the silence period should be greater than the T2 timer defined in ITU-T Rec. T.30.
  - MGC signalling or other out of band signalling method.
- In the direction from IP to GSTN network due to receipt of RTP packets that have non-VBD payload types, only after the first VBD RTP packet has been received. This will avoid the situation of an incorrect transition into Audio mode when it has transitioned to VBD mode on detection of VBD signals on its TDM side and is still receiving Voice RTP packets (because the remote end has not yet transitioned, based on reception of the VBD RTP packets).

The above described transition criteria are also summarized in Figure 1.

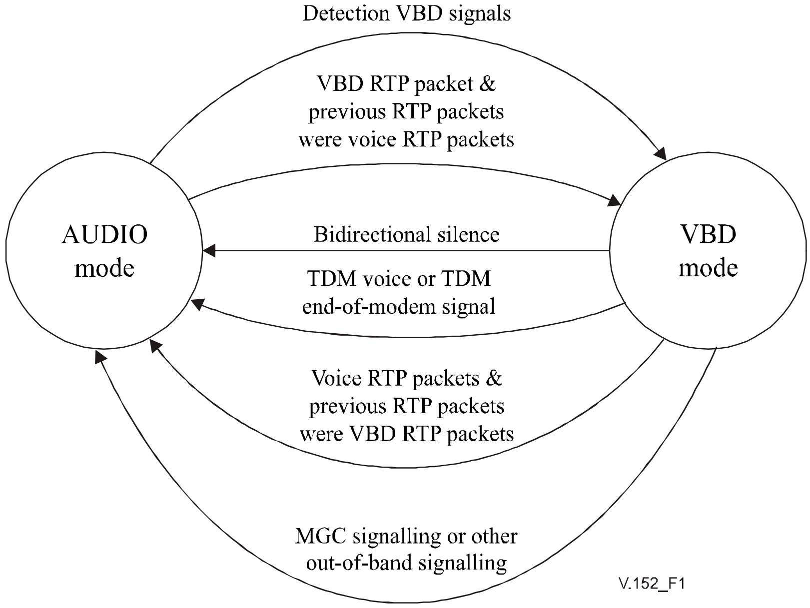

Figure 1/V.152 – Voice-VBD transitioning state diagram

11 Optional procedures for indicating to a remote end transition into VBD using State Signalling Events (SSEs)

This clause describes the procedures that a V.152 implementation must use when using the State Signalling Event (SSE) protocol as defined in Annexes C, E and F of V.150.1.

Note that the use of SSE is optional for V.152 implementation and subject to negotiation with the remote gateway. When one or both of the two gateways does not support SSE operation, then transitions to and from the VBD mode shall be governed by the procedures defined in clause 10.

11.1 Declaration of SSEs

The SSE capability shall be signalled as defined in F.6/V.150.1. The minimum set of state signalling events that shall be supported for VBD operation are events 0 through 3 which are basic to the SSE protocol. The null SSE event (0) shall never be sent and should be ignored if received.

11.2 Transition to the VBD mode for V.150.1 gateways

When both gateways support V.150.1, then transitions to and from the VBD mode is governed by V.150.1 procedures. These transitions are synchronized via the SSE protocol.

11.3 Transition to the VBD mode for non-V.150 cases

When one or both of the two gateways does not support V.150.1 operation, then transitions to and from the VBD mode shall be governed by the procedures in this clause. An attempt shall be made to make these procedures isomorphic to V.150.1 procedures so that media gateways do not incur the burden of supporting and testing multiple VBD switching mechanisms.

A V.152-compliant media gateway that has successfully negotiated support of SSE media gateway shall respond to a voice-band data stimulus by immediately transitioning the connection to the VBD media state and issuing an SSE indicating this state (C.5.2/V.150.1). As with all other media state transitions, this is contingent on resource availability. On making this transition locally, the stimulus-detecting media gateway may start sending VBD packets immediately.

On receiving an SSE indicating the VBD media state (SSE:VBD), a media gateway shall immediately transition the connection to the VBD media state if it has the resources to do so. Before making this transition, it may ignore any in-band VBD packets it receives (20.4/V.150.1).

The transition to a VBD media state in response to detecting a voice-band data stimulus (such as an answer tone variant) is illustrated in Figure 2. In this example, the on-ramp (call-originating) gateway G1 and the off-ramp (call-terminating) gateway G2 support VBD operation.

On detecting a voice-band data stimulus, gateway G2 determines whether it has the resources to transition the session to the VBD media state. If it does, then it immediately makes the transition and sends SSE:VBD (event code 2) to gateway G1. While in the VBD media state, it uses an RTP payload type marked for VBD treatment.

On receiving SSE:VBD, gateway G1 determines whether it has the resources to transition the session to the VBD media state. If it does, then it immediately makes the transition and sends an SSE:VBD back to gateway G2, confirming that its media state has changed to VBD. If it does not, then it sends an SSE:audio (event code 1) on receiving a SSE:VBD from G2.

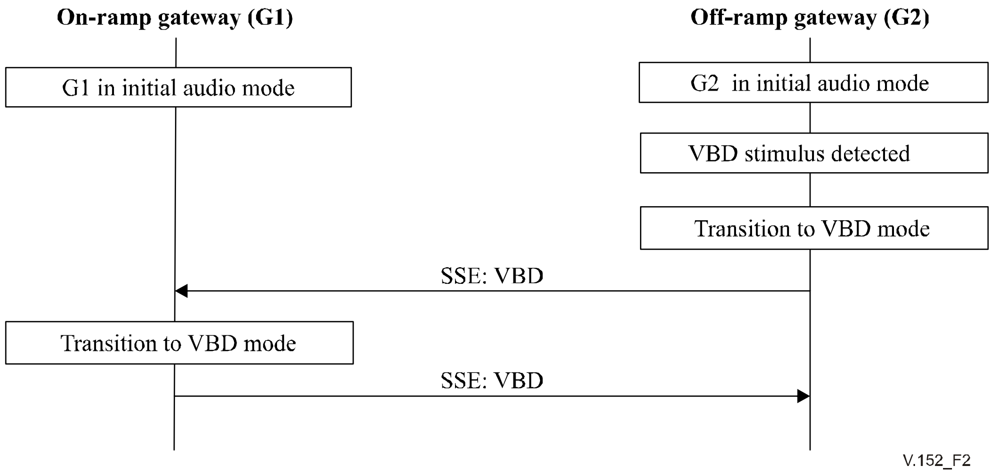

Figure 2/V.152 – Initiation of VBD operation in response to VBD stimulus detection

When sending SSE:VBD, gateways G1 and G2 may use suitable Reason Information Codes (RICs) defined in ITU-T Rec. V.150.1. An example is the RIC indicating an answer tone. A null code, which conveys no information, may also be used. The SSE:VBD from G1 may indicate a p' state transition as the Reason Information Code. Since p' is defined as a gateway's view of the other gateway's protocol state, this indicates that this SSE from G1 is a response to a received SSE.

Distinctions based on RIC codes (V.8 versus non-V.8, text vs. non-text) may be used to optimize play-out buffer settings and FEC levels for different applications of VBD. Additionally, when the RIC indicates text, gateways may optimize operation by doing one of the following:

- Loading both the audio and VBD codecs to streamline quick, on-the-fly transitions between talkspurts and textspurts.
- Staying in the VBD mode across talkspurts and textspurts.

11.4 Transition from the VBD media mode

On detecting a termination of data transmission, media gateways shall locally transition the connection to the audio mode and issue an SSE:audio (Initial audio SSE, event code 1) to the remote gateway. If it receives an SSE:audio, it shall change the media state to initial audio and respond back with an SSE:audio.

The criteria for determining the termination of data transmission are application-specific and are not defined here. Examples of such criteria are the detection of voice or of pre-determined intervals of silence. A transition to the modem, fax or text media states is not a termination of data transmission.

By declaring support of the SSE protocol, gateways implicitly declare support for events 1-3 which are basic to the protocol. In order to be used, support for other SSEs such as SSE:FR (Facsimile relay SSE, event code 4) and SSE:TR (Text relay SSE, event code 5) shall be explicitly declared.

Transitions from the VBD media state to the MR (modem relay), FR (facsimile relay) and TR (text relay) states are permitted. These media state transitions are contingent on:

1) Declaration of the applicable capabilities at call establishment.
2) Availability of resources at the time of the state transition.

In the modem relay case, these are synchronized using SSE:MR (Modem relay SSE, event code 3) per ITU-T Rec. V.150.1.

SSE:TR (Text relay SSE, event code 5) is recommended to synchronize media shifts from VBD to the text relay mode, and vice versa. For example, when V.21 signals are followed by Annex A/V.18 signals in end-to-end V.18 automoding, there may be a shift to VBD based on the ANS preceding V.21, and possibly another shift to TR if the gateway does not support VBD for Annex A/V.18.

For autonomous media shifts from VBD to facsimile relay, two cases arise from T.38:

1) For gateways that comply with V.150.1 and Annex F/T.38, SSE:FR is used. Both single-port and multi-port operation is supported.
2) For all other gateways, port activity monitoring is used. Note that single-port operation for audio RTP and T.38 udptl packets is not supported; however, single-port operation may be used if audio RTP and the optional T.38 RTP procedure are being used.

Since Facsimile switchovers are tolerant in terms of signal timing, external signalling can be used in lieu of the autonomous VBD-to-FR media shifts described in the last paragraph. Examples of external signalling are SIP re-invites, H.245 RequestMode/CLC/OLC and H.248.1 context modification. End-to-end timing issues that often jeopardize the use of external signalling with modem traffic do not exist for Facsimile traffic.

For a session that is in the VBD media state, a gateway may reject SSE:MR, SSE:FR or SSE:TR with an SSE:VBD or SSE:audio. If used, SSE:audio causes a transitioning of the session to an Audio mode.

The SSE Reason Identifier codes for VBD mode are defined in Table 12/V.150.1.

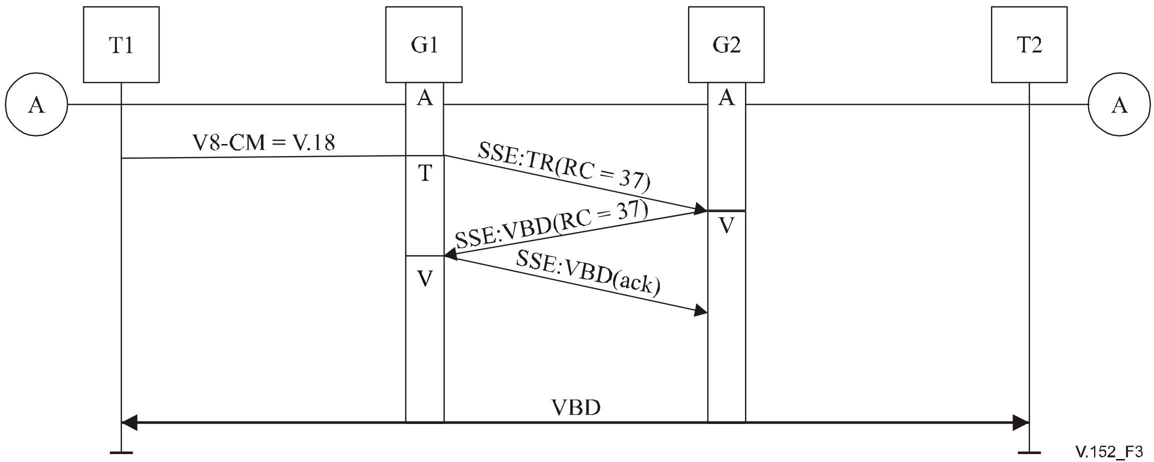

Figure 3/V.152 – V.18 text telephone using VBD

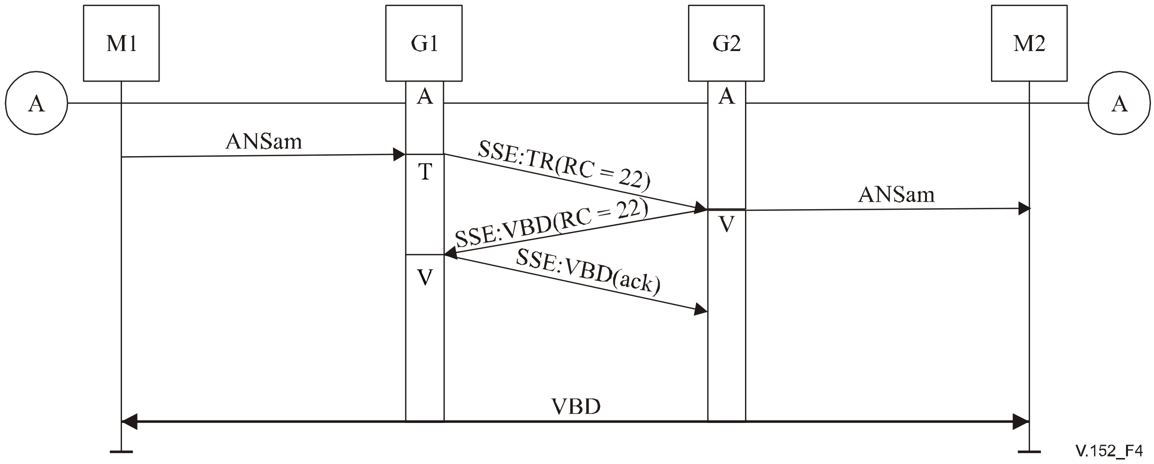

Figure 4/V.152 – V.34 modem using VBD

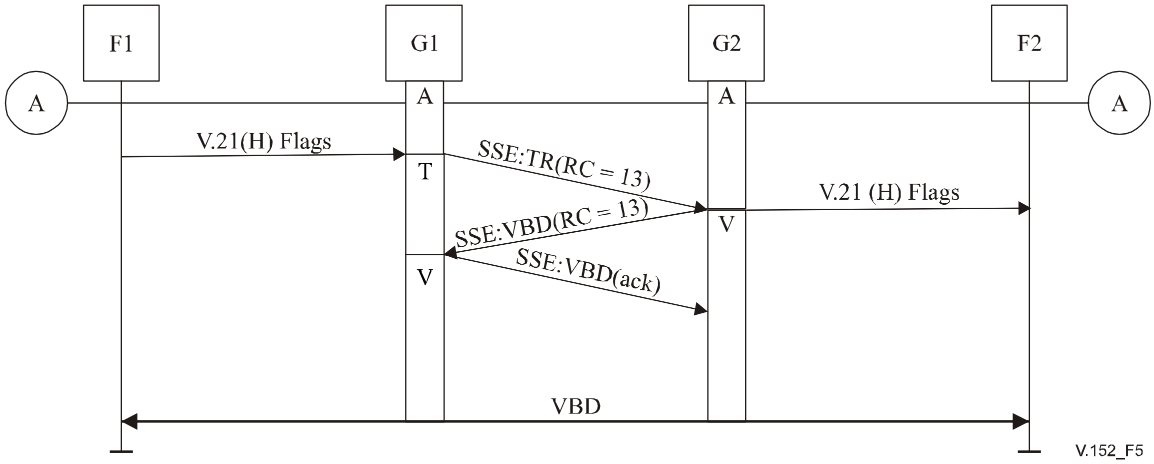

Figure 5/V.152 – Group 3 facsimile (without CNG/CED tones) using VBD

11.5 Security – Optional

When the VBD mode is used to transport data payloads, it lends itself easily to secure, encrypted operation based on SRTP (Secure RTP). Support of security features is not required by V.152 compliant implementation and is negotiated at call establishment.

Based on declarations at the time of session establishment, it is possible to encrypt some RTP payload types (e.g., voice, VBD and RFC 2833 events), while passing other RTP payload types (e.g., SSEs) unencrypted. Such selective encryption will allow fast response times to SSEs, without compromising the security of media sourced by the end-user. When one end proposes encrypted operation for a set of payload formats, and the other end does not support encryption, the preferred outcome is a rejection of the proposal and termination of the connection attempt. At this point, either the on-ramp or off-ramp media gateway or media gateway controller can counter-propose a non-encrypted connection through the call signalling protocol(s) in use.

Annex A

Vendor-defined messages

Vendor-specific messages may be supported within V.152, subject to negotiation with the remote end. In general, a V.152 implementation may support up to 255 vendor identifiers (vendor-ID) for a given call. Each vendor-ID may be unique or specific and tied to either a single or a multiple set of attributes. A unique vendor-Tag may also be assigned to each set of attributes associated with a vendor-ID to allow simpler use within V.152.

Usually the vendor-ID is provided during the external signalling used during the call set-up (i.e., H.245, H.248 or SDP, etc.). The format used in signalling schemes may be compliant to either ITU-T Rec. T.35 or the IANA private enterprise number. The choice is up to the vendor.

When the vendor-ID format is ITU-T Rec. T.35, the vendor-ID consists of a country code followed by a vendor code. The country code consists of four octets and the vendor-ID consists of two octets. If the representation of the vendor-ID is hexadecimal, leading zeros in the country code may be omitted, while leading zeros in the vendor code may not be omitted.

When the vendor-ID is the vendor's IANA private enterprise number, leading zeros may be omitted.

The vendor-Tag is a decimal integer with a value between 0 and 255. If used, values in the range of 1 to 255 are uniquely mapped to the combination of vendor-ID and vendor-specific information. The choice of this integer made by a gateway is independent of the choice made by its peer gateway. Due to the compactness of this index, a gateway or endpoint may use it in a number of places to simplify the messaging. A value of zero for the vendor-Tag is a null value. When present, it is equivalent to omitting the vendor-Tag. A null value of the vendor-Tag is not associated with any vendor-ID. If non-null, the vendor-Tag may serve as a dynamically assigned vendor-specific identifier.

The vendor-specific information is an octet string consisting of one or more octets as defined by the vendor. Since it consists of an integer number of octets, it is represented by an even number of hex characters. No "0x" prefix is needed. Limitation on size is context specific. Details where size is limited will be indicated appropriately.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	**Data communication over the telephone network**
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2005