International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.1020

(07/2006)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Quality of service and performance – Generic and user-related aspects

---

Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks

ITU-T Recommendation G.1020

ITU-T

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TRANSMISSION MEDIA CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
QUALITY OF SERVICE AND PERFORMANCE – GENERIC AND USER-RELATED ASPECTS	**G.1000–G.1999**
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DATA OVER TRANSPORT – GENERIC ASPECTS	G.7000–G.7999
PACKET OVER TRANSPORT ASPECTS	G.8000–G.8999
ACCESS NETWORKS	G.9000–G.9999

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

ITU-T Recommendation G.1020

Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks

Summary

The transmission of speech and other voiceband applications over packet networks brings with it new, sometimes unique, forms of quality degradation. There are many existing definitions for packet network performance parameters, yet the desire to control the quality of non-elastic isochronous applications requires additional, complementary information. The purpose of this Recommendation is to define packet network and terminal performance parameters that better reflect the perceived quality of the target applications. It is largely focused on quality impairments resulting from delay variation and packet loss which are peculiar to IP and other packet-based technologies, and which do not appear in traditional TDM networks. It discusses the interactions and trade-offs among these packet impairments, and describes mechanisms such as de-jitter buffers and packet loss concealment for reducing their effects on the quality of speech and other applications. However, this Recommendation avoids overlap by making reference to existing definitions wherever possible.

The parameters defined by this Recommendation extend beyond the IP layer in many cases. End-to-end packet system (combination of end terminals and network) parameters are also necessary to determine the speech/voiceband quality. Clauses 5, 6 and 7 collect the parameter definitions for source terminals, packet networks, and destination terminals, respectively. Overall parameters are defined in clause 8.

In this version, the Recommendation adds formal definitions for many of the voice metrics identified in RFC 3611, RTCP Extended Reports (RTCP-XR).

Annex A defines the gateway-specific reference points and parameters. Annex B provides information on packet loss distributions and packet loss models, and defines burst packet loss parameters. Annex C specifies an adaptive de-jitter buffer emulator and defines de-jitter buffer parameters. Appendix I tabulates the RTCP-XR parameters, along with the examination existing Recommendations and the final action, which typically was the addition of a parameter in this Recommendation.

Source

ITU-T Recommendation G.1020 was approved on 14 July 2006 by ITU-T Study Group 12 (2005-2008) under the ITU-T Recommendation A.8 procedure.

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 not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

CONTENTS

Page

Introduction

The transmission of speech and other voiceband applications over packet networks brings with it new, sometimes unique, forms of quality degradation. There are many existing definitions for packet network performance parameters, yet the desire to control the quality of non-elastic isochronous applications requires additional, complementary information. The purpose of this Recommendation is to define packet network and terminal performance parameters that better reflect the perceived quality of the target applications, extending beyond the IP layer in many cases. End-to-end packet system (combination of end terminals and network) parameters are also necessary to determine the speech/voiceband quality, and this Recommendation defines them as well.

ITU-T Recommendation G.1020

Performance parameter definitions for quality of speech and other voiceband applications utilizing IP networks

1 Scope

This Recommendation defines a set of performance parameters for packet networks and end terminals that can assist in quantifying the end-to-end quality of speech and other voiceband applications. It is largely focused on quality impairments resulting from delay variation and packet loss which are peculiar to IP and other packet-based technologies, and which do not appear in traditional TDM networks. It discusses the interactions and trade-offs among these packet impairments, and describes mechanisms such as de-jitter buffers and packet loss concealment for reducing their effects on the quality of speech and other applications.

This Recommendation recognizes existing performance parameter definitions, and avoids duplication. Many factors that determine the quality of speech and voiceband applications are common to both TDM and IP-based networks, and are addressed in existing Recommendations. In the parlance of ITU-T Rec. I.350, the scope of this Recommendation is limited to the information transfer function of the 3×3 matrix, and only to the bearer channel. Call processing aspects of connection access and disengagement (e.g., dialtone delay and post-dialling delay) are not considered in this Recommendation. Furthermore, this Recommendation does not specify numerical objectives for packet networks or end terminals, although this will be the subject of follow-on work.

Figure 1 illustrates this scope, along with some other specifications with their areas of coverage. This Recommendation only defines parameters that describe packet terminal and packet transmission impairments that are unique to speech and voiceband application quality assessment.

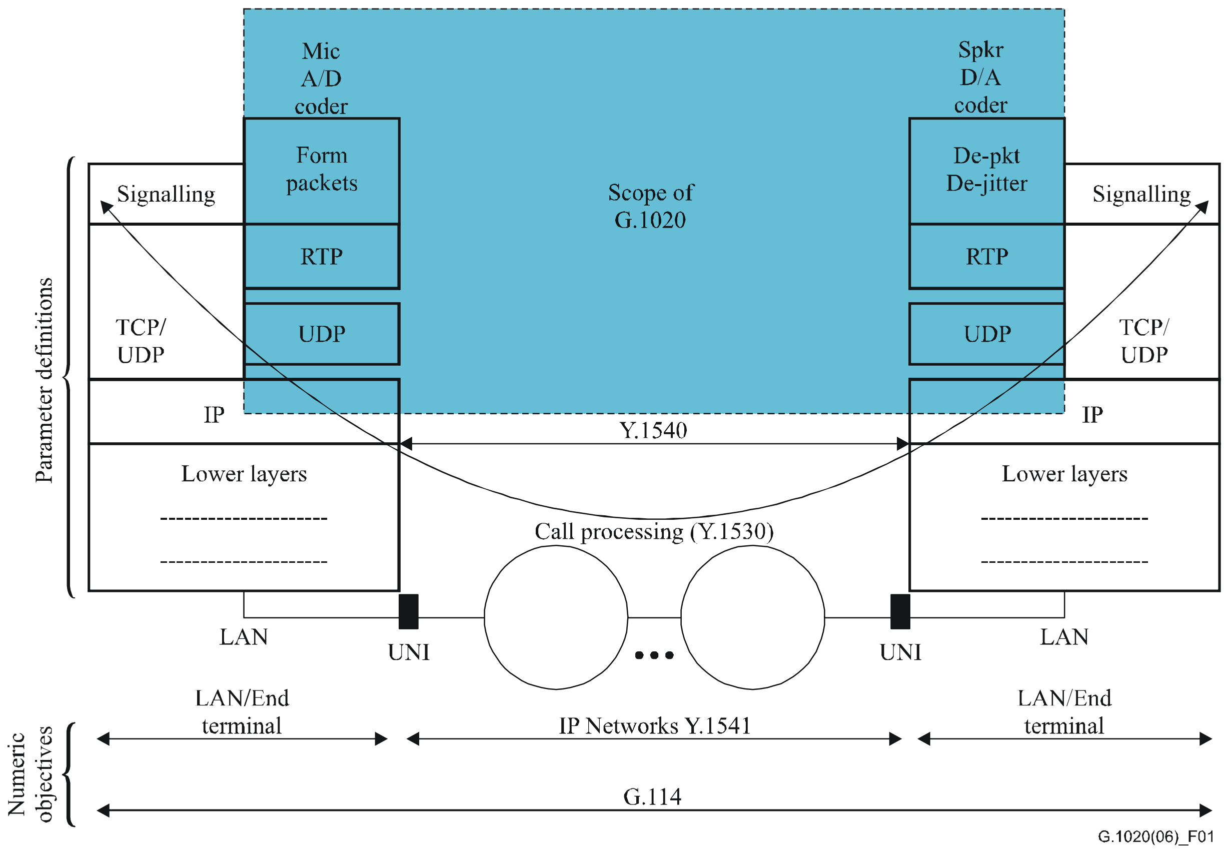

Figure 1/G.1020 – Scope of ITU-T Rec. G.1020 as it relates to other performance specifications

Note that the number of IP networks between terminals is not limited in these definitions.

Other ITU-T Recommendations supplement the parameters provided in this Recommendation. For example, transmission planning over hybrid Internet/PSTN connections is covered in ITU-T Rec. G.177. Still, other Recommendations specify such parameters in the context of assessing the performance of IP packet transfer on international data communication services (e.g., ITU-T Rec. Y.1540). Network performance objectives for different QoS classes of IP-based services are described in ITU-T Rec. Y.1541, and end-to-end one-way delay objectives are specified in ITU-T Rec. G.114.

New Recommendations that complement ITU-T Rec. G.1020 are anticipated. Call processing aspects of the connection are in development. The ITU-T is also currently working on new Recommendations dealing with situations where the network is mainly an IP network with islands of PSTN, and where the network is mainly a PSTN network with islands of IP. ITU-T work to specify the performance objectives for VoIP terminals and gateways was recently completed in ITU-T Rec. P.1010. Finally, there is work on methods to assess the performance by measuring metrics related to the end-to-end quality of VoIP.

The definitions of packet transmission parameters that are unique to ATM networks are explicitly out of the scope of this Recommendation.

This Recommendation should be particularly useful to those new to the area of Voice over IP (VoIP) who want to gain a better understanding of the factors affecting the quality of these telecommunication systems. Developers of telecommunication equipment can use the parameters defined in this Recommendation to specify relevant aspects of their contribution to end-to-end performance. Service providers can use these parameters to effectively summarize performance of IP network solutions.

2 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 regularly published. 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.107 (2005), *The E-model, a computational model for use in transmission planning*.

– ITU-T Recommendation G.113 (2001), *Transmission impairments due to speech processing*.

– ITU-T Recommendation G.114 (2003), *One-way transmission time*.

– ITU-T Recommendation G.177 (1999), *Transmission planning for voiceband services over hybrid Internet/PSTN connections*.

– ITU-T Recommendation I.356 (2000), *B-ISDN ATM layer cell transfer performance*.

– ITU-T Recommendation P.51 (1996), *Artificial mouth*.

– ITU-T Recommendation P.57 (2005), *Artificial ears*.

– ITU-T Recommendation P.1010 (2004), *Fundamental voice transmission objectives for VoIP terminals and gateways*.

– ITU-T Recommendation Y.1540 (2002), *Internet protocol data communication service – IP packet transfer and availability performance parameters*.

– ITU-T Recommendation Y.1541 (2006), *Network performance objectives for IP-based services*.

– IETF RFC 3611 (2003), *RTP Control Protocol Extended Reports (RTCP XR)*.

3 Definitions

This Recommendation defines the following terms:

3.1 ear reference point: A virtual point for geometric reference located at the entrance to the listener's ear, traditionally used for calculating telephonometric loudness ratings [P.57].

3.2 terminal input measurement point: A measurement point in the physical medium connecting a terminal to an IP network that is crossed as IP packets leave the IP network and enter the terminal. This measurement point is as close to the terminal as possible.

3.3 IP terminal: An endpoint device intended for connecting to an IP network to support speech communications. These devices can be dedicated (e.g., a telephone set) or general purpose (e.g., a computer running an application that performs the terminal function).

3.4 terminal output reference point: A measurement point in the physical medium connecting a terminal to an IP network that is crossed as IP packets leave the terminal and enter the IP network. This measurement point is as close to the terminal as possible.

3.5 de-jitter buffer: A buffer designed to remove the delay variation (i.e., jitter) in packet arrival times. Data is put into the de-jitter buffer at a variable rate (i.e., whenever they are received from the network), and taken out at a constant rate.

3.6 mouth reference point: The point on the reference axis, 25 mm in front of the lip plane [P.51].

3.7 real-time signal: A signal accurately representing acoustic or electrical signals in the time domain.

3.8 receive electrical reference point: An electrical point of reference that is equivalent to the ear reference point from a terminal delay measurement perspective.

3.9 send electrical reference point: An electrical point of reference that is equivalent to the mouth reference point from a terminal delay measurement perspective.

4 Abbreviations

This Recommendation uses the following abbreviations:

ADC	Analogue-to-Digital Converter
ATM	Asynchronous Transfer Mode
DAC	Digital-to-Analogue Converter
DSCP	Differentiated Services Code Point
Dst	Destination
HDLC	High-level Data Link Control
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPER	IP Packet Error Ratio
IPErr	Errored IP Packet Count
IPLR	IP Packet Loss Ratio
IPPM	IP Performance Metrics working group
IPRE	IP packet transfer Reference Event
IPSLB	IP Packet Severe Loss Block
IPSLBR	IP Packet Severe Loss Block Ratio
IPTD	IP Packet Transfer Delay
IPv4	Internet Protocol version 4
IPv6	Internet Protocol version 6
MAPDV2	Mean Absolute Packet Delay Variation 2
NA	Not Available
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RFC	Request For Comments
RSVP-TE	Resource Reservation Protocol – Traffic Engineering
RTCP	Real-time Transport Control Protocol
RTP	Real-time Transport Protocol
RTPErr	Errored RTP Packet Count
SPR	Spurious Packet Ratio
Src	Source

TDM	Time Division Multiplex
UDP	User Datagram Protocol
UDPErr	Errored UDP Packet Count
UNI	User-Network Interface

5 Source terminal packet parameters

This clause gives the relevant sending terminal packet parameters that have a direct effect on perceived speech and voiceband application quality. Figure 2 indicates the positions of measurement points and system components.

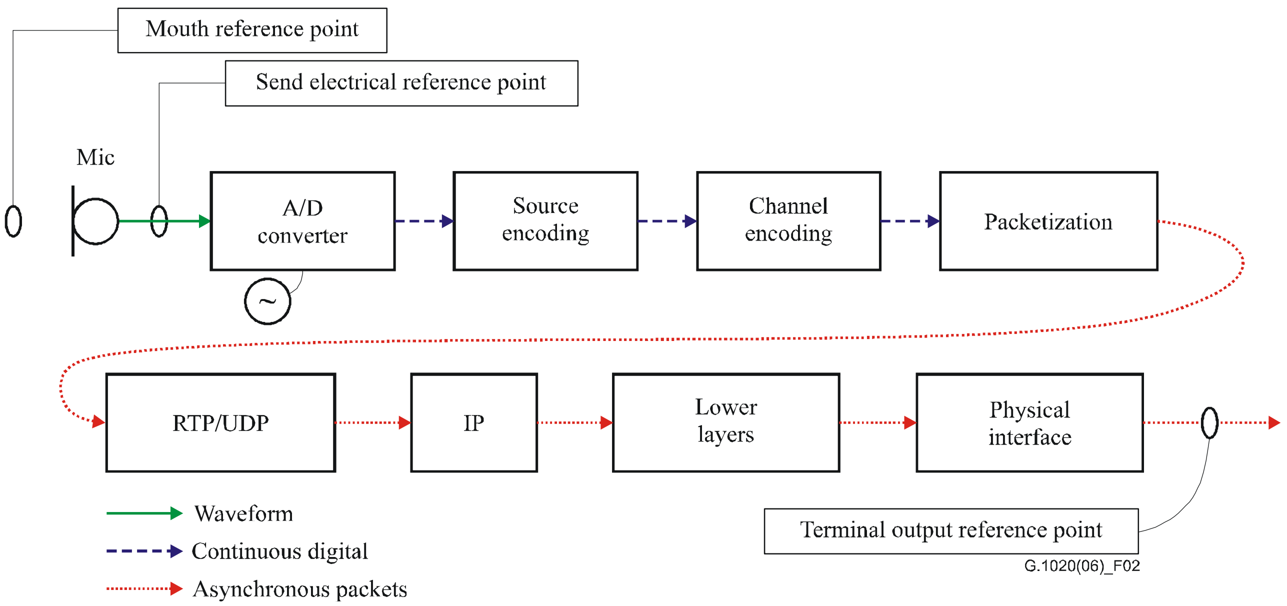

Figure 2/G.1020 – Source terminal diagram and reference points

5.1 Analogue/digital conversion clock accuracy (free-running)

The relative frequency offset of a clock may be specified as:

$$\frac{\Delta f}{f_{\text{nominal}}} = \frac{f_{\text{measured}} - f_{\text{nominal}}}{f_{\text{nominal}}}$$

The measured frequency error should be minimized, and use of an atomic frequency reference for measurement is preferred (the nominal frequency should be as close to ideal as practical).

The long-term (or end-of-life) accuracy of the clock oscillator technology (e.g., quartz crystal) may also be given if known, as it establishes an upper bound on the frequency offset.

For example, assume that the nominal frequency of the analogue/digital conversion oscillator is 8000 Hz. A measurement on the source terminal oscillator indicates a frequency of 8000.0027 Hz. The relative frequency offset is then 0.0027/8000 = 3.38×10^{-7}. Disciplined quartz oscillators can usually achieve offsets of 1×10^{-6} based on their accuracy specifications.

Note that it should be possible to infer the sending clock frequency from measurements of the source packet rate (under favourable circumstances, for example, when silence suppression is disabled). This would permit a measurement using externally available signals. The method is for further study.

Relative frequency offset is a unit-less quantity, often specified as a fraction or in parts per million.

5.2 Packet information field size

The Packet information field size specifies the amount of encoded voiceband waveform that a packet contains. This size must be expressed unambiguously, using as many of the following units of measure as necessary:

1) The number of 8-bit octets of encoded voiceband signals and supporting information (e.g., bit associated with forward error correction to aid packet loss concealment, or bits associated with encryption).
2) The number of encoder frames (the specific encoder and native frame size must also be specified).
3) The amount of continuous waveform time represented by the encoded bits in the field.

Typical information fields contain one or two encoder frames and combine 10 or 20 ms of waveform time in a single packet. Annex A/G.114 provides guidance on the calculation of delay for various coders and packet information field sizes.

5.3 Packet overhead

The total octets appended to the packet information field should be counted separately for each protocol layer header. The total octets dedicated to non-information headers may be counted as the packet exits the terminal output reference point, therefore, including the effects of header compression, if any. The octet size of packets dedicated to media flow control, status/performance reports (e.g., RTCP), or other non-media carrying packets should be counted separately. Some typical overhead contributors are (in octets):

RTP:12; UDP:8; IPv4:20; IPv6:40; HDLC Encapsulation:8; Flag:1.

The unit of measure for packet overhead is 8-bit octets.

5.4 Source terminal delay

The source terminal delay is the interval defined by the time that a signal enters the mouth reference point and the time that the first bit of the corresponding encoded, packetized signal exits the terminal output reference point. When appropriate, the send electrical reference point may be substituted for the mouth reference point. By definition, the source terminal delay includes the entire packetization/depacketization time, and test waveforms and methods must retain sufficient information to assess packetization time variability due to alignment between the test signal and the packet boundaries. For example, the test signal must be of sufficient length to span packet boundaries, permitting boundary identification in time. The portion of the signal that is carried by the earliest part of the packet payload should be used.

NOTE – This delay will include source terminal delay variation if present, and appropriate statistics should be applied to summarize the variation.

The unit of measure for source terminal delay is time in seconds.

5.5 Source terminal delay variation

The fundamental notion of a 1-point delay variation parameter is the comparison between the actual packet emission pattern and the intended (usually periodic) emission pattern. Some variations of this definition include a "skipping clock" adjustment, as in ITU-T Rec. I.356.

The source terminal delay variation is defined as the time difference between the first bit of a packet emission at the terminal output reference point and the ideal periodic reference time. For the first packet in a flow, the ideal periodic reference time is set equal to the emission time. Subsequent packets emissions are compared to this periodic time reference, as shown below:

$$\text{Source Terminal Delay Variation} = t(packet\_n) - t(reference\_packet\_n)$$

where *t(reference_packet_n)* is the emission time of *packet_n* of the ideal periodic reference stream. The time interval of measurement, along with appropriate statistics, should be provided.

NOTE – Long interval measurements may include the undesired effects of source frequency offset. Variation due to source frequency offset should be noted and removed as a measurement error, when possible.

The unit of measure for source terminal delay variation is seconds.

6 Packet network performance parameters

Standards for IP layer packet transfer performance (see e.g., ITU-T Recs Y.1540, Y.1541 and IETF RFCs 2330, 2678 through 2681, 3357 and 3393) include the following parameters: One-way transfer delay, delay variation, and packet loss. These parameters must be mapped to the application layer to adequately estimate user impact.

6.1 Summary of network performance parameters

Table 1 summarizes the IP network performance parameters relevant to this Recommendation, and the relationships among network and terminal parameters that form the basis for overall system parameters. Reading left to right, each row identifies a parameter and indicates how it may be combined with other parameters to derive a specific overall performance parameter for one aspect of the end-to-end or user-user quality (although the exact formulas are given in later clauses).

Table 1/G.1020 – Summary of IP network performance parameters and mapping to overall/user-user parameters

IP network parameter (Note)	Translation to overall	Overall parameter
Transfer delay (IPTD, mean)	IPTD + Src delay + Dst delay	Mean user-user delay
Delay variation (IPDV, 99.9%tile minus the minimum)	Combine with Src delay variation distribution	Contributes to Dst delay, or audio frame loss
Delay jump (possibly captured by RFC 3393, for example)	May come from network path/facility change, or may only appear at de-jitter buffer output	Audio time-scale discontinuity
Errored packet (headers)	IPErr + UDPErr + RTPErr ⇒	Audio frame loss (packet or codec frame discard)
Reordered packet (Appendix VII/Y.1540)	(may be considered lost) ⇒	Audio frame loss
Lost packet	IP loss + (all audio defects) ⇒	Audio frame loss (pre-concealment)
IP severe loss block (IPSLB)	(depends on block duration)	Call cut-off
Loss patterns (e.g., RFC 3357)	Complete stream loss/arrival ⇒	Burst length/Consecutive loss
Packet rate (inferred from other system characteristics)	Difference of source and destination terminal ADC and DAC oscillators	System frequency offset (relative to destination)
NOTE – From ITU-T Rec. Y.1540, unless noted.		

6.2 Additional network parameters recommended

All the fundamental single packet outcomes are defined in ITU-T Rec. Y.1540 (and IETF RFCs listed in Bibliography). However, it is possible to derive additional parameters of interest when streams or flows of packets are considered, as in VoIP planning and measurement.

6.2.1 Consecutive packet loss event

For cases where successive packets, sent in a periodic stream (according to RFC 3432, for example), are designated as lost according to the definition of lost packet outcome in ITU-T Rec. Y.1540, then the length of the event should be specified as the number of packets lost in sequence. This length should be recorded separately for each event. Following a measurement encountering multiple consecutive loss events, the count for each event length should also be recorded. Sequence numbers contained in packet headers may assist this measurement.

6.2.2 Degraded second

A degraded second outcome occurs for a block of packets observed during a 1-second interval when the ratio of lost packets at the egress UNI to total packets in the corresponding second interval at the ingress UNI exceeds D%. Sequence numbers and timestamps contained in packet headers may be used to aid in this measurement.

The value of D is provisionally set at 15%, and may change on the basis of further experience or study. For example, if a flow of packets at 50 packets per second is impaired by 8 losses (16%), then the quality will be degraded whether the losses are consecutive or distributed throughout the second.

The unit of measure for the degraded second is a count of these outcomes.

6.2.3 Short-term IP delay variation/jitter quantification

The following clauses provide two approaches to short-term jitter quantification. When the distribution of delays over short intervals is available, then the first approach, based on short-term range is recommended. However, if the complete time series of delay variation is known, then the approach based on mean absolute packet delay variation may provide additional information.

6.2.3.1 Approach based on short-term range

This definition is consistent with Appendix II/Y.1541.

Short-term IP delay variation is defined as the maximum IPTD minus the minimum IPTD during a given short measurement interval.

$$IPDV_{Short\_Term} = IPTD_{max} - IPTD_{min}$$

where:

$IPTD_{max}$ is the maximum IPTD recorded during the short measurement interval;

$IPTD_{min}$ is the minimum IPTD recorded during the short measurement interval.

This is a simple and fairly accurate method for calculating IPDV in real-time. The length of the short measurement interval is for further study. The measurement interval influences the ability of the metric to capture low and high frequency variations in the IP packet delay behaviour.

To be consistent with other parameter definitions in this Recommendation, a measurement interval of 1 second is provisionally agreed.

Many values of $IPDV_{Short\_Term}$ are measured over a longer time interval (comprising many short measurement intervals). The 99.9th percentile of these $IPDV_{Short\_Term}$ values is expected to meet the Y.1541 objective of 50 ms (note that this objective was established for a 1-minute measurement interval and the percentile is evaluated on a per-packet basis, assuming a 50-packet per-second sending rate or higher).

As an example, assume 1200 one-second measurements of $IPDV_{Short\_Term}$, collected over 20 minutes. If two or more measurements of $IPDV_{Short\_Term}$ exceed 50 ms, then the Y.1541 objective may not have been met during a few intervals, and a more exact evaluation of the objective is warranted.

The unit of measure for short-term range is seconds.

6.2.3.2 Approach based on mean absolute packet delay variation

An alternative approach is to determine the mean absolute packet delay variation with regard to a short-term average or minimum value – termed here the *adjusted absolute packet delay variation*. This may provide a more meaningful relationship to de-jitter buffer behaviour.

For packet 1: (or any packet arriving after 3 consecutive lost packets have been detected)

$D_1 = t_1$

Deviation $t_1 – D_1 = 0$

should not be counted as Positive or Negative

Do not compute MAPDV2 for the first packet.

For subsequent packets:

mean delay $D_i = (15 \times D_{i–1} + t_{i–1})/16$

if $t_i > D_i$ then

positive deviation $P_i = (7\times P_{i–1} + t_i – D_i)/8$

negative deviation $N_i = (7 \times N_{i–1})/8$

else (when $t_i \le D_i$)

positive deviation $P_i = (7 \times P_{i–1})/8$

negative deviation $N_i = (7 \times N_{i–1} + D_i – t_i)/8$

NOTE – For the first packet, t_1, $t_{i–1} = 0$ and $D_{i–1} = 0$, so $t_1 = D_1$, and deviations are not computed.

We compute mean absolute packet delay variation 2 (*MAPDV2*) for packet (*i*) as:

$$MAPDV2 = (P_i) + (N_i)$$

Note that the number of consecutive lost packets needed to reset the algorithm is based on an assumption that three lost packets will cause a typical de-jitter buffer to empty, and require priming as done at the beginning of a stream of packets.

The unit of measure for MAPDV2 is seconds.

7 Destination terminal packet parameters

This clause gives the relevant destination terminal packet parameters that have a direct effect on perceived speech and voiceband application quality, and a set of overall packet parameters. Figure 3 indicates the positions of measurement points and system components.

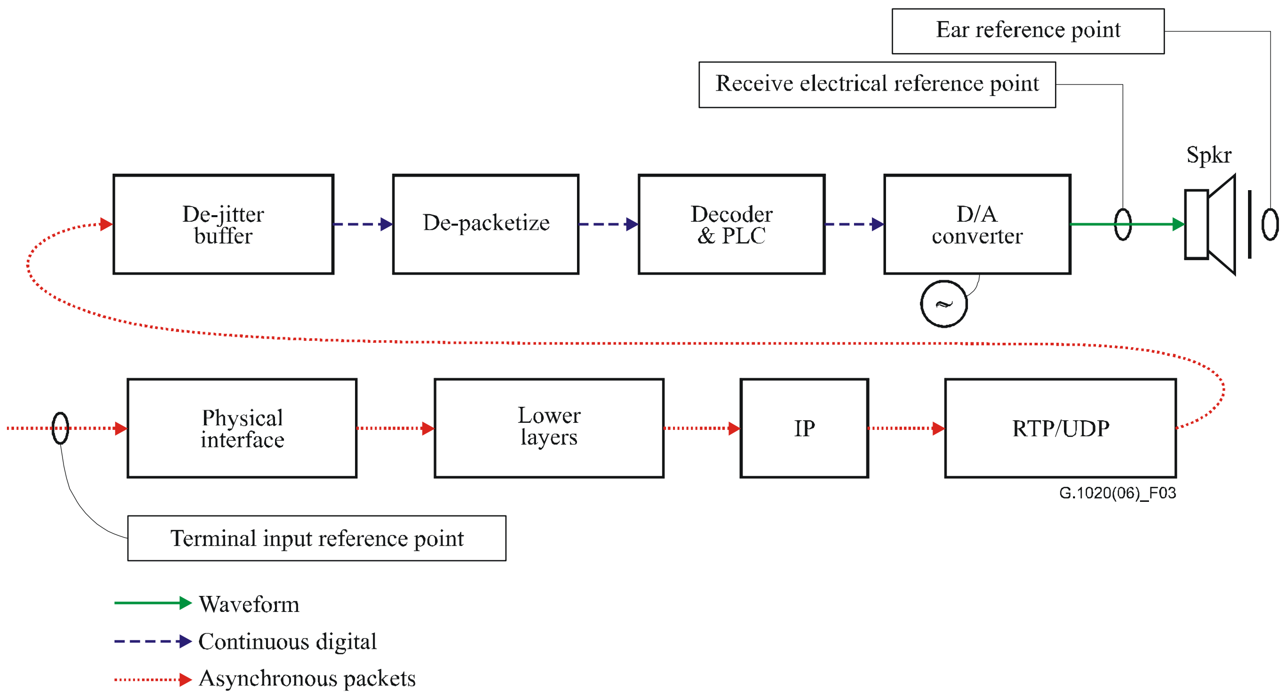

Figure 3/G.1020 – Destination terminal components

7.1 Discussion of destination packet processing

Figure 4 depicts the process through which IP packet parameters/impairments (transfer delay, delay variation, and packet loss and errors) can be mapped to application layer performance in terms of overall loss and delay.

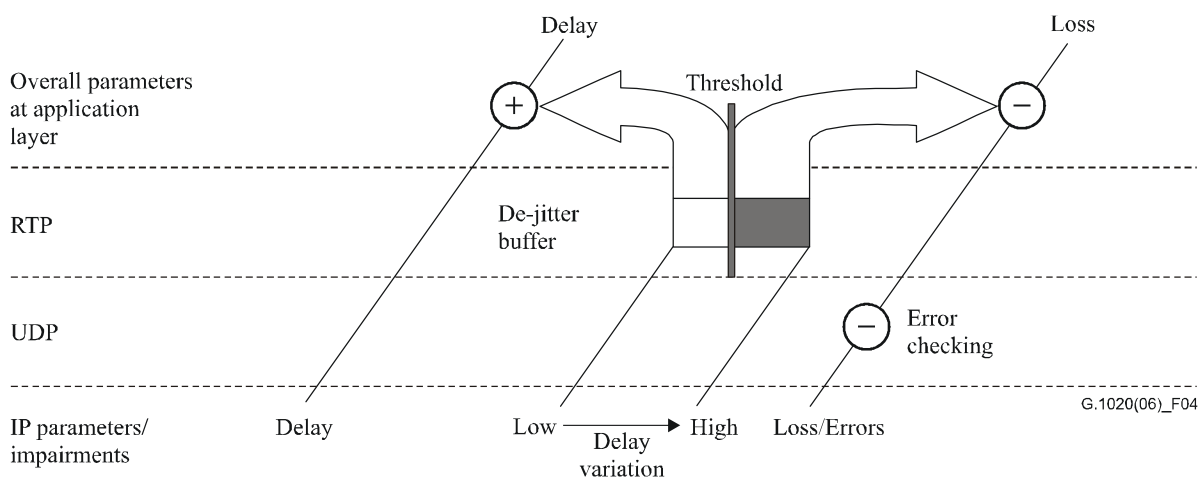

Figure 4/G.1020 – Mapping IP packet performance to application layer

At the bottom of the figure, packets arrive with various impairments due to the source terminal and network(s), or never arrive (lost). The arriving packets are processed as they move up the protocol stack to remove as much impairment as possible. We show that some forms of impairment (errors, jitter) map into other impairments (overall loss, overall delay).

Figure 4 captures the trade-off between application level delay and loss as a threshold on the range of delay variation based on the size of the de-jitter buffer. Packets with delay variation in the "white" range are accommodated, while packets with larger variation (in the "black" range) would be discarded. A larger de-jitter buffer can accommodate packets with greater delay variation, hence, fewer packets would be lost overall at the expense of larger overall delay. Conversely, a smaller de-jitter buffer will produce less overall delay, but expose a larger fraction of packets to be discarded by the terminal and increase the overall loss.

7.2 Taxonomy of de-jitter buffer types/parameters and models

There are two main types of de-jitter buffers, fixed length and adaptive length. De-jitter buffers can be constructed in many different ways, including the following attributes identified in Table 2. The values of applicable de-jitter buffer parameters must be known when assessing the performance of a system.

Table 2/G.1020 – De-jitter buffer types and parameters

Type	Attributes	Possibilities	
Fixed (and adaptive)	Size (configure maximum and nominal or minimum)	Integer number of packets	Fractional number of packets
Adaptive	Control	Timed delay if no over/under flow	Evaluate loss ratio (configure lowest acceptable threshold, and minimum packet count between adjustments)
	Adjustment	Timed	Silence gaps only
	Initialization	First packet	Small sample
	Adjustment granularity	Packet size	Fraction of packet size
	Restores packet order	Yes	No
	Voiceband data mode	Detect 2100 Hz tone, set to maximum length	None

7.2.1 Destination terminal delay and loss assessment

The primary contributors to delay are variable sources. This clause illustrates how the de-jitter buffer size and network IP packet delay variation overlap, and how one must carefully accumulate specific delay statistics to achieve the correct delay totals.

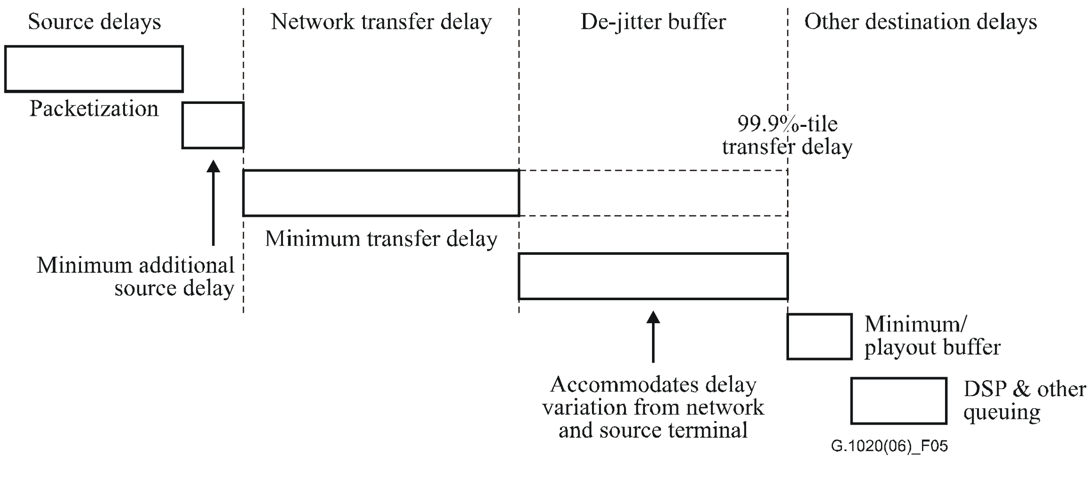

Figure 5/G.1020 – Delay of packet networks and network elements

Figure 5 shows some of the elements of a VoIP path that contribute to end-end delay. At the sender, packetization time can be significant. There is usually a variable delay as packets traverse the network. At the receiver, the de-jitter buffer exists to accommodate the delay variation and deliver a continuous payload stream. We note that packets with the minimum source and network delay spend the maximum time in the de-jitter buffer; likewise packets that encounter delays longer than the minimum spend less time in the buffer. There is also some minimum amount of time each packet must spend in a buffer at the receiver; possibly as long as an entire packet.

The following subclauses give an overview of the process to combine the IP layer loss and delay with the additional contributions from destination terminal higher-layer functions, such as the de-jitter buffer.

7.2.1.1 Loss

Depending on the type of de-jitter buffer, there will be some criteria for determining whether or not each specific packet in a flow is accommodated or discarded. The result can completely change the distribution of overall packet losses. For example, if random bit errors are causing packets to fail the UDP checksum, then packet losses will have a random distribution as they proceed to the application layer. But, if several consecutive packets experience excessive delays, then the additional discards due to the limitations of the de-jitter buffer will make the overall loss distribution appear bursty rather than random. Therefore, categorization of the loss distribution must take place at the application layer (using techniques such as in Annex B, or the burst ratio, see Appendix I/G.113), before estimation of application performance with tools such as the E-model (see ITU-T Rec. G.107).

There are circumstances where packet order may change during network transfer. Some de-jitter buffers are unable to restore order to these reordered packets (ITU-T Rec. Y.1540) and, in this case, they are designated as discarded packets.

7.2.1.2 Delay

The correct value of buffer delay to combine with the other delays depends on the descriptive statistics available. For example, the mean network delay should be summed with the average de-jitter buffer occupation time (and other delays) to obtain an overall average delay. This method allows for buffer adaptation, needing only the average queuing time for all packets in the assessment time interval. On the other hand, if only the minimum network delay is known, it should be summed with the maximum de-jitter buffer occupation (or size used, and other delays) to give an overall delay.

We next consider initialization for a fixed-size de-jitter buffer. If the first packet to arrive has the minimum transfer delay, then the receiver will buffer the packet for the entire time requested, and buffering size will be as expected. Fortunately, many packets arrive at or near the minimum transfer time, so this case has a fair likelihood.

On the other hand, if the first packet has a rather long delay, then more buffer space will be needed to accommodate the "early" arrival of packets at or near the minimum transfer time, and the de-jitter buffer will contribute more than the expected delay to the overall calculation.

7.2.1.3 Fixed de-jitter buffer and destination terminal model

The simplest effective model of loss due to a fixed de-jitter buffer is to designate as discarded all packets whose delay is greater than the minimum transfer delay for the packet stream plus the (fixed) de-jitter buffer length.

The following procedure provides a mapping between the IP and application layers, assuming fixed length de-jitter buffers for destination terminal performance assessment.

1) Designate as lost all packets failing the UDP Checksum.

2) Designate as discarded all packets whose delay is greater than the minimum transfer delay for the packet stream plus the (fixed) de-jitter buffer length, or whose delay is less than the established minimum.

3) Sum the mean network delay (IPTD) with the average source and destination terminal delay to obtain an overall average delay, OR, sum the minimum source terminal delay and the minimum network delay with the maximum destination terminal delay (reflecting the maximum de-jitter buffer occupation when network jitter is present, or maximum size used).

In step 2 above, the minimum transfer delay should be evaluated over short intervals (provisionally a value of 10 seconds is used). The minimum for the first interval is used throughout, unless the short term minimum grows beyond the accommodation range of the buffer. In this case, no packets will be delivered to the upper layers and the de-jitter buffer must be reset to the new minimum, as would likely occur in practice. Alternatively, if the short term minimum should fall to a value where a high percentage (provisionally 50%) of packets would be designated lost due to early arrival, the de-jitter buffer must be reset to the new minimum.

When calculating the overall impairment contribution of a fixed de-jitter buffer, the distribution of delay variation determines the proportion of packets that would be discarded. The distribution of packet delays that are accommodated (not discarded) can be used to calculate the mean de-jitter buffer occupation delay, as follows:

Mean occupation delay = [De-jitter Buffer Size] – (Mean Delay of Accom. Packets – Min. Delay)

This mean delay can be added with other destination terminal delay constants to produce an estimate of the mean destination terminal delay. If the exact delay distribution is not available, then there is agreement that a value of half the de-jitter buffer size may be substituted for the mean occupation delay in calculations supporting network planning.

If the maximum destination terminal delay is needed in calculations, then the maximum de-jitter buffer size can be added with other destination terminal delay constants to produce an estimate of the maximum delay.

7.2.1.4 Adaptive de-jitter buffer model

The fixed de-jitter buffer in item 2 above may be replaced with an adaptive de-jitter buffer emulation, as described in this clause when a time series of packet stream information is at hand.

The time series of packet arrivals may be used with an adaptive de-jitter buffer emulator to determine the buffer size dynamics and the mean de-jitter buffer occupation time (delay) over the series. This mean delay can be combined with other destination terminal delay constants to produce an estimate of the mean destination terminal delay.

An example of an adaptive de-jitter buffer emulator is provided in Annex C.

7.2.2 Destination terminal loss-related parameters

This clause defines two key parameters used in RTCP-XR [RFC 3611] reports, loss rate and discard rate, and several supporting terms. The definitions of these parameters are somewhat unique because they must be based solely on information present at the receiver, while most definitions of packet loss include sender knowledge and receiver observation. The principle ambiguity is avoided by declaring loss only when there is certainty that there is a gap in the received sequence numbers (e.g., when a packet with a larger than expected number arrives), as this is evidence that the sender did not simply stop sending.

7.2.2.1 Packet play-out time

The play-out time for any packet is determined by adding a fixed amount of time to the RTP timestamp. This added time includes the time for the packet to traverse the network and any buffering applied at the receiver, so that the play-out time is the instant the packet is removed from the buffer for decoding or other processing before presentation. Sequence numbers and scheduled play-out times are paired and tracked throughout the life of the stream.

7.2.2.2 Discard window

The discard window is the time interval for acceptable packet arrival referenced to the play-out time, possibly having asymmetric tolerance for early or late arrival.

7.2.2.3 Discarded packets

When a packet arrives and its play-out time is calculated (by adding a fixed time to the timestamp), the packet is declared discarded if the limits of the discard window have been exceeded.

7.2.2.4 Maximum waiting time for lost packets

The maximum waiting time for any packet is determined by adding a fixed amount of time to the play-out time estimated for that packet. This time will typically be significantly longer than the discard window as applied to late packets.

7.2.2.5 Lost packets

When the play-out time for a given packet is past by more than the maximum waiting time, and a packet with a larger sequence number has arrived, then the packet is declared lost.

7.2.2.6 Dimensions for evaluation

In the parameters above, time is expressed in seconds. It would be possible to evaluate these same parameters using the dimensions of packets, providing that each packet represents a fixed amount of play-out time (as is usually the case in VoIP). For example, the discard window would be expressed in packets using their sequence numbers, and if a packet arrives having a sequence number outside the bounds of the discard window, it is discarded. The maximum waiting time would instead be expressed as a number of packets with respect to the sequence number of the current packet being played-out.

7.2.2.7 Loss ratio for use in the RTCP-XR loss rate parameter

The ratio of the total number of packets declared lost to the total number expected from the beginning of reception to the reporting time, where the total number expected is determined from the earliest and latest sequence numbers.

The unit of this parameter is a ratio of lost to total packets.

7.2.2.8 Discard ratio for use in the RTCP-XR discard rate parameter

The ratio of the total number of packets declared discarded to the total number expected from the beginning of reception to the reporting time, where the total number expected is determined from the earliest and latest sequence numbers.

The unit of this parameter is a ratio of discarded to total packets.

7.2.3 Destination terminal delay

The destination terminal delay is the interval defined as beginning when the first bit of a packet representing a waveform signal enters the terminal input reference point and ending when the corresponding decoded, depacketized signal exits the ear reference point. When appropriate, the receive electrical reference point may be substituted for the ear reference point.

NOTE – This delay may vary if an adaptive de-jitter buffer is present, and appropriate statistics should be applied to summarize the variation.

Since, by definition, the source terminal delay includes the entire packetization/depacketization time, destination terminal packet test signals should be constructed such that they occupy the earliest part of the payload. In this way, source and destination terminal delay measurements will be conducted at equivalent moments with respect to packetization time.

The unit of measure for destination terminal delay is time in seconds.

7.3 System frequency offset, using destination clock as reference

The system frequency offset may be assessed by monitoring sequence number increment per unit time or accumulated time-stamp offset, and is a measure of the difference between source and destination analogue/digital conversion clock accuracy. The relative frequency offset between the source and destination clocks may be specified as:

$$\frac{\Delta f}{f_{Destination}} = \frac{f_{Source} - f_{Destination}}{f_{Destination}}$$

This frequency offset may be used to determine the rate of buffer overflow or underflow events at the destination terminal, usually resulting in additional packet losses, by noting that the fractional frequency offset is equivalent to the time shift (Δt) over an observation interval (T).

$$\frac{\Delta f}{f_{Destination}} = -\frac{\Delta t}{T}$$

(noting that frequency and time period differences have a negative relationship). For example, assume that the source frequency is 7999.997 Hz, the destination frequency is 8000.001 Hz, and the de-jitter buffer length is 20 ms. Since the destination's D/A converter clock is reading information faster than the source supplies it, the de-jitter buffer will eventually empty, or underflow. At a relative offset of:

$$\frac{7999.997 - 8000.001}{8000.001} = -5 \times 10^{-7}$$

(where the minus sign indicates that the source clock pulses occur slower than the corresponding pulses at the destination), the time shift equal to the entire de-jitter buffer will accumulate in an observation interval of:

$$T = \frac{-(\Delta t = 0.02)}{-5 \times 10^{-7}} = 40\,000\,\mathrm{sec} = 667\,\mathrm{min}$$

System frequency offset is a unit-less quantity, often specified as a fraction or in parts per million.

7.4 Packet loss concealment (type, delay)

Many standardized speech coders have a native Packet Loss Concealment (PLC), and it is sufficient to specify whether or not the PLC is on or off, and account for any additional delay. For example, Appendix I/G.711 adds at least 3.75 ms algorithmic delay, and possibly more depending on implementation. This PLC may be used with other waveform coders, such as G.726.

Many forms of non-standardized PLC have emerged in practice, particularly for G.711 and other waveform coders. If these are used, the specific PLC algorithm and delay should be specified.

Note that a PLC that sounds best to human users may not meet the needs of voiceband modem carrier detectors. If there is a signal classifier for voiceband data or fax modems, and a special PLC is selected to improve their operation on packet networks, then the PLC type and signal classification method should be specified.

8 Overall performance parameters

8.1 Overall delay (including source, network and destination)

Following the analysis of the de-jitter buffer and other destination terminal components, as described throughout 7.2.1 and illustrated in Figure 5, it is possible to combine them with the delay of the source terminal and network(s) to determine the system's overall delay. The following formulas are acceptable, and their use is determined by the specific delay statistics available for computation.

When the mean delays for all components are at hand:

Overall_Mean_Delay = mean(source_delay) + mean(net_delay) + mean(destination_delay)

As Figure 5 clearly illustrates, the minimum delays for source terminal and network can be combined with the maximum destination terminal delay to obtain an estimate using constants:

Overall Delay (constant) = min(source_delay) + min(net_delay) + max(destination_delay)

When measured directly, the overall delay is the interval defined from the time that a signal enters the mouth reference point and ending at the time when the corresponding signal exits the ear reference point (or equivalent reference points).

One method for direct measurement of overall delay has been documented in Annex B of [10].

Several examples of the overall mean delay calculation are present in Appendix VII/Y.1541. These examples utilize various network configurations and reference terminals with several packet sizes, de-jitter buffers, and forms of packet loss concealment. Appendix VII/Y.1541 goes on to calculate E-model R values for each of these cases [G.107].

The unit of measure for overall delay is seconds.

8.2 End System Delay

This clause defines the End System Delay parameter used in RTCP-XR [RFC 3611] reports.

The end system delay is defined as the sum of the Source delay and Destination delay for the end system reporting the value, and includes the sample accumulation delay, encoding delay, de-jitter buffer delay, decoding delay, and play out delay. The value may be estimated or measured.

The unit of measure for end system delay is seconds.

8.3 Round Trip Delay

This clause defines the Round Trip Delay parameter used in RTCP-XR [RFC 3611] reports.

The round trip time is defined as the time a packet takes to travel from the source RTP interface to the destination RTP interface and return to the source RTP interface. The most recently calculated time is reported.

The unit of measure for round trip delay is seconds.

8.4 Time-scale discontinuities in post-de-jitter and PLC stream

A time-scale discontinuity is defined as a sudden change in the overall delay, measured from the mouth reference point to the ear reference point. This parameter captures how often the user's time reference shifts, due to the network path, the de-jitter buffer, or both.

The units of measure for time-scale discontinuities are seconds, and an event count.

8.5 Overall (frame/packet) loss (including network and destination)

This parameter may be expressed in terms of packets, or coder frames. It is important to understand the relationship between frame loss and packet loss. For example, when two frames are combined in each packet, then every packet lost implies a burst of two frame losses and the decoder/PLC must attempt recovery under these more difficult circumstances than the isolated single frame loss.

8.5.1 Overall (frame/packet) loss ratio

The overall loss ratio for an evaluation interval is defined as follows:

$$\text{Overall Loss Ratio} =$$

$$1 - \frac{(\text{Total\_pkt\_sent} - \text{Lost\_net} - \text{Lost\_error\_check} - \text{Discarded\_de-jitter-buffer} - \text{Discarded\_reordering})}{\text{Total\_pkt\_sent}}$$

The unit of this parameter is a ratio of lost to total packets (or frames).

8.5.2 Overall (frame/packet) loss model

In order to assess the impact of losses and discards on VoIP applications, it is useful to consider distribution of these impairments over time. Typical approaches include the Gilbert-Elliott model and similar Markov-based models. RFC 3611 specifies the use of a Gilbert-Elliott model to describe packet loss and discard distribution and gives an example of a four-state Markov model to derive these parameters. Annex B provides a description of these models, and gives an example of a typical packet loss/discard distribution. The typical output parameters are the average gap length and loss/discard density, and the average burst length and loss/discard density.

8.5.3 Overall consecutive (frame/packet) loss event count

After examining a stream of packets sent according to [9], and a set of successive packets have been designated as lost or discarded according to all the relevant loss/discard criteria in the overall loss ratio parameter, then the length of the event should be specified as the number of packets lost in sequence. This length should be specified separately for each event, in units of packets. The count of each event size should also be provided as a result. Sequence numbers contained in packet headers may be used to aid in this measurement.

8.5.4 Pitfalls and errors in calculating overall parameters

One simplified approach to obtain the overall end-to-end delay has been to take the mean IP packet transfer delay, and combine it with constants for other elements in the mouth-to-ear path. This procedure may produce errors because of variable delays in some terminal components (e.g., the de-jitter buffer), or because the variable delay elements are ignored.

Another potential pitfall would be to use the packet loss ratio as measured by a test receiver that allows, for example, 3 seconds before declaring a packet lost, thereby underestimating the loss ratio. A typical de-jitter buffer would have much less tolerance for long delays beyond the norm. Hereto, knowledge of the de-jitter buffer figures prominently in the mapping between IP packet performance and loss at the application layer.

Annex A

VoIP gateway-specific reference points and performance parameters

A.1 Introduction

VoIP gateways are usually deployed to interconnect packet and circuit switched networks and require new reference points for delay and other parameters. This annex defines the gateway-specific reference points and parameters.

A.2 Definitions

This annex defines the following terms:

A.2.1 packet input reference point: A measurement point in the physical medium connecting an IP network to a gateway that is crossed as IP packets leave the IP network and enter the gateway. This measurement point is as close to the terminal as possible.

A.2.2 packet output reference point: A measurement point in the physical medium connecting a gateway to an IP network that is crossed as IP packets leave the gateway and enter the IP network. This measurement point is as close to the gateway as possible.

A.2.3 TDM input reference point: A measurement point in the physical medium connecting a Time Division Multiplex Network to a VoIP gateway. Signals that cross this point are packetized and enter the IP network. This measurement point is as close to the gateway as possible.

A.2.4 TDM output reference point: A measurement point in the physical medium connecting a gateway to a Time Division Multiplex Network. Signals that cross this point are carried to the end terminal. This measurement point is as close to the gateway as possible.

A.3 Source gateway parameters

This clause gives the relevant source gateway packet parameters that have a direct effect on perceived speech and voiceband application quality. Figure A.1 indicates the positions of measurement points and system components.

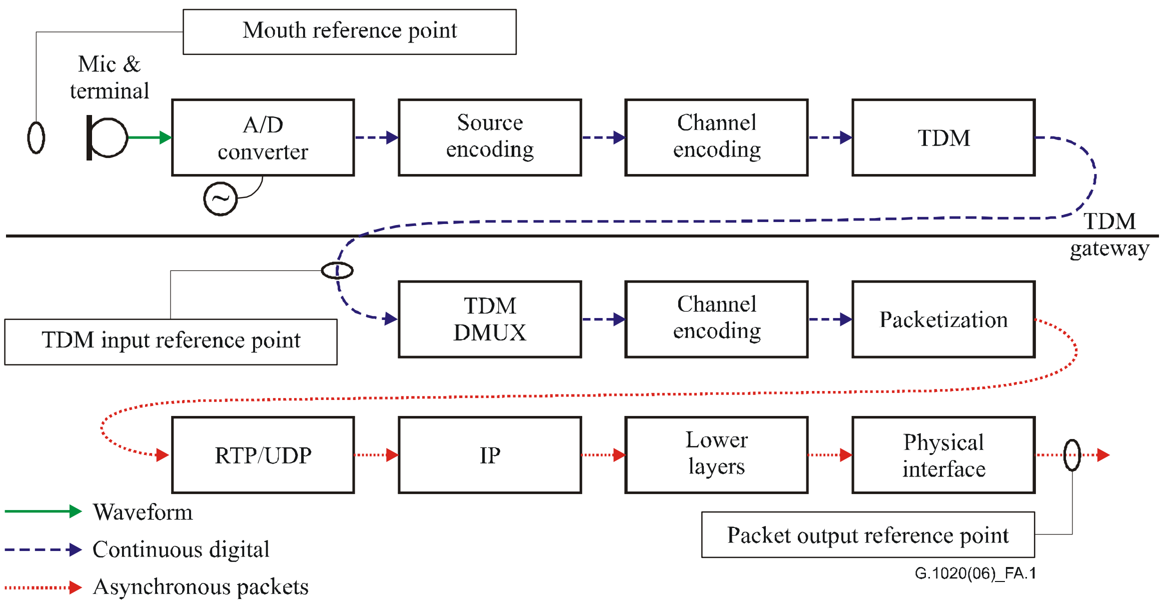

Figure A.1/G.1020 – Source gateway diagram and reference points

We note that some gateways will include coders that provide bit rate compression, while others will simply packetize the PCM waveform, or provide other processing to the voiceband signals such as Fax Demod/Remod.

Most of the source terminal parameters defined in clause 5 are relevant to source gateways. Some parameters require the following reference point substitutions in Table A.1.

Table A.1/G.1020 – Source Reference Point Substitution

Source Terminal	Substitute Source Gateway Reference Point
Mouth Reference Point	TDM Input Reference Point
Send Electrical Reference Point	TDM Input Reference Point
Terminal Output Reference Point	Packet Output Reference Point

The source terminal to source gateway parameter mapping is as follows:

Table A.2/G.1020 – Source Parameter Mapping

Source Terminal Parameter	Source Gateway Parameter
Source Terminal Delay	Source Gateway Delay
Source Terminal Delay Variation	Source Gateway Delay Variation

Parameters such as Packet Information Field Size require no modifications.

A.4 Destination gateway parameters

This clause gives the relevant destination gateway packet parameters that have a direct effect on perceived speech and voiceband application quality. Figure A.2 indicates the positions of measurement points and system components.

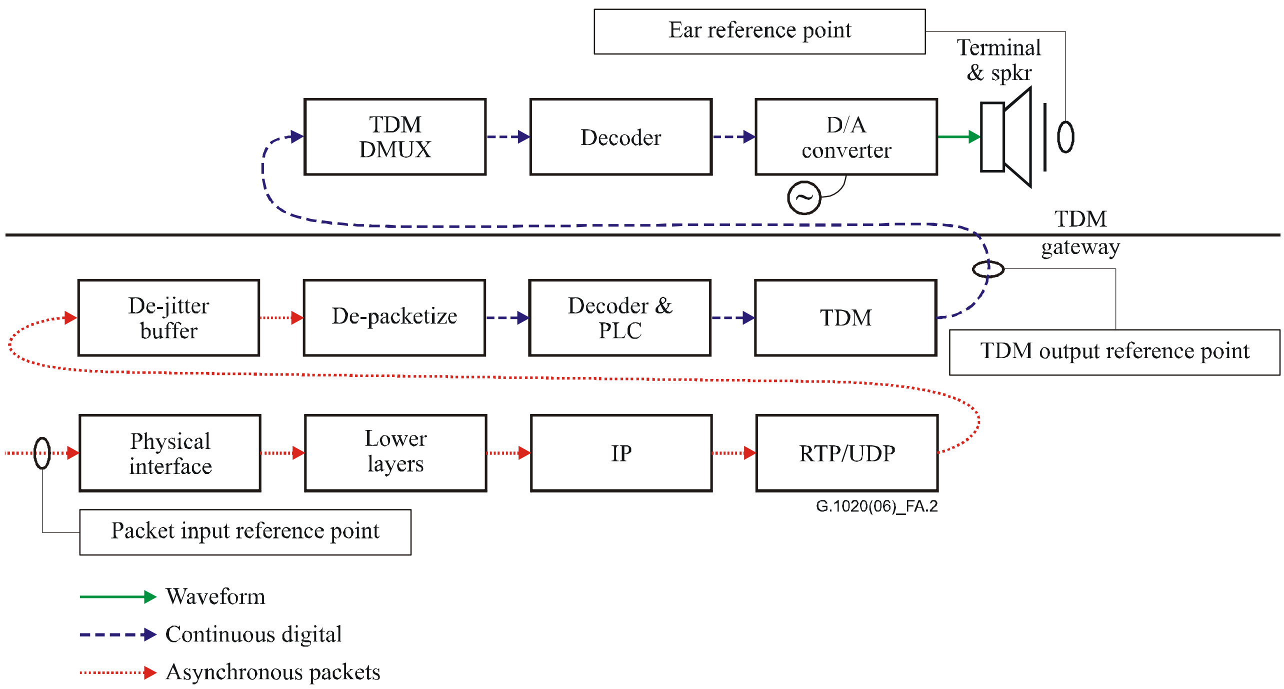

Figure A.2/G.1020 – Destination gateway diagram and reference points

Most of the destination terminal parameters defined in clause 7 are relevant to destination gateways. Some parameters require the reference point substitutions in Table A.3.

Table A.3/G.1020 – Destination Reference Point Substitution

Destination Terminal	Substitute Destination Gateway Reference Point
Ear Reference Point	TDM Output Reference Point
Receive Electrical Reference Point	TDM Output Reference Point
Terminal Input Reference Point	Packet Input Reference Point

The destination terminal to destination gateway parameter mapping is as follows.

Table A.4/G.1020 – Destination Parameter Mapping

Destination Terminal Parameter	Destination Gateway Parameter
Destination Terminal Delay	Destination Gateway Delay

Parameters such as Packet Loss Concealment require no modifications.

A.5 Overall delay

When a gateway is present in the end-end path, the additional delay in the TDM network between the Mouth or Ear Reference point and the gateway must be included in the overall delay. The TDM network elements usually have fixed delays, so they can be added to the delay for the packet network components. The unit of measure for overall delay is seconds.

Annex B

Packet loss distributions and packet loss models

B.1 Introduction

It is generally understood that packet loss distribution in IP networks is "bursty", however, there is less certainty concerning the use of specific loss models and, in fact, some misunderstanding related to some commonly used models, for example the Gilbert Model. This annex outlines some key packet loss models, provides some analysis of packet loss data, and discusses the degree of "fit" of models and data.

B.2 Common packet loss models

B.2.1 Historical background

Much of the early work on loss or error modelling occurred in the 1960's in relation to the distribution of bit errors on telephone channels.

One approach used was a Markov or multi-state model. Gilbert [B-13] appears to be the first to describe a burst error model of this type, later extended by Elliott [B-10] and [B-11] and Cain and Simpson [B-6]. Blank and Trafton [B-3] produced higher state Markov models to represent error distributions.

Another approach was to identify the statistical distribution of gaps. Mertz [B-16] used hyperbolic distributions and Berger and Mandelbrot [B-2] used Pareto distributions to model inter-error gaps. Lewis and Cox [B-15] found that in measured error distributions there was strong positive correlation between adjacent gaps.

Packet loss modelling in IP networks seems to have followed a similar course, although the root cause of loss (typically congestion) may be different to that of bit errors (typically circuit noise or jitter).

B.2.2 Bernoulli or independent model

The most widely used model is a simple independent loss channel, in which a packet is lost (or bit error occurs) with a probability P_e. For some large number of packets N, then the expected number of lost packets is $N \times P_e$. The loss probability can be estimated by counting the number of lost packets and dividing this by the total number of packets transmitted.

B.2.3 Gilbert and Gilbert-Elliott models

The most widely known burst model is the Gilbert model [B-13] and a variant known as the Gilbert-Elliott model [B-10] and [B-11]. These are both two state models that transition between a "good" or gap state 0 and a "bad" or burst state 1 according to state transition probabilities P_{01} and P_{11}:

i) *Gilbert model*;
   a) State 0 is a zero loss/error state;
   b) State 1 is a lossy state with independent loss probability P_{e1};

ii) *Gilbert-Elliott model*;
   a) State 0 is a low loss state with independent loss probability P_{e0};
   b) State 1 is a lossy state with independent loss probability P_{e1}.

It is often assumed that the Gilbert model lossy state corresponds to a "loss" state, i.e., that the probability of packet loss in state 1 is 1, however, this is incorrect (it would be more proper to describe this as a 2-state Markov model). This leads to analysis of packet loss burstiness in terms solely of consecutive loss which misses the effects of longer periods of high loss density. These long periods of high loss density can have significant effect on voice over IP services.

For example, consider the following:

Loss pattern 00000110010101011011000000000000000000

Correct application of Gilbert Model – burst length 15, burst density 60%

Incorrect application of Gilbert Model – mean burst length 1.5 packets

B.2.4 Markov models

A Markov model is a general multi-state model in which a system switches between states i and j with some transition probability $p(i, j)$.

A 2-state Markov model has some merit in that it is able to capture very short-term dependencies between lost packets, i.e., consecutive losses [B-1], [B-4], [B-14] and [B-18]. These are generally very short duration events (say 1-3 packets in length) but occasional link failures can result in very long loss sequences extending to tens of seconds [B-5].

By combining the 2-state model with a Gilbert-Elliott model, it is possible to capture both very short duration consecutive loss events and longer lower density events.

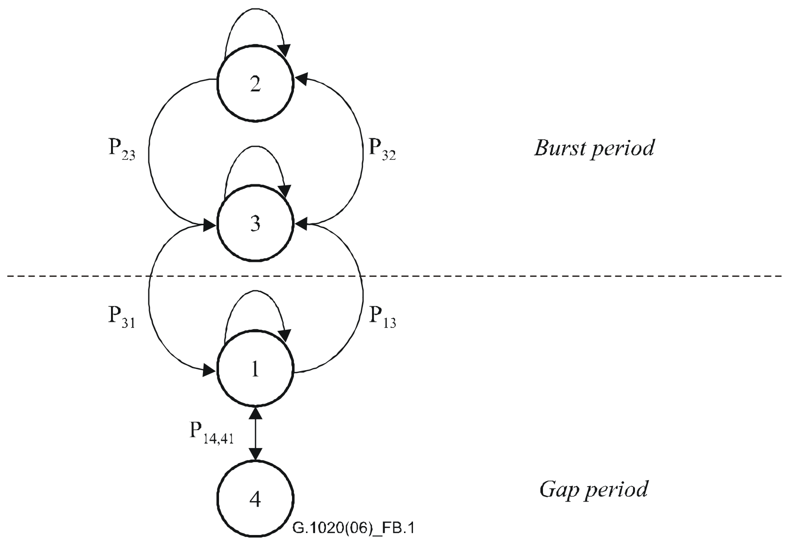

Figure B.1/G.1020 – 4-state Markov model

This 4-state Markov model [B-7] and [B-12] represents burst periods, during which packets are received and lost according to a first 2-state model and gap periods during which packets are received and lost according to a second 2-state model. The states have the following definition:

a) State 1 – packet received successfully;

b) State 2 – packet received within a burst;

c) State 3 – packet lost within a burst;

d) State 4 – isolated packet lost within a gap.

For example, using the loss pattern below:

Loss pattern	00000110010101011011000000000000000000000001000000000
State	11111332232323233233111111111111111111111114111111111

It is common to define a gap state with respect to some criteria, for example, a loss rate lower than some limit or some consecutive number of received packets. A convenient definition is that a burst must be a longest sequence beginning and ending with a loss during which the number of consecutive received packets is less than some value G_{min} (a suitable value for G_{min} for use with Voice over IP services would be 16 whereas, for use with video services, a higher value of say 64 or 128 would be preferable). These definitions are given more formally below.

B.2.5 Burst parameters

This clause defines the Burst Parameters used in RTCP-XR [RFC 3611] reports.

B.2.5.1 G_{min} parameter

The G_{min} parameter setting gives the number of consecutive received packets that distinguish whether the stream is experiencing a burst, or is in the gap between bursts. The units are packets.

B.2.5.2 Burst

A burst is the longest sequence of packets beginning and ending with a loss, where the sequences of consecutive received packets are less than G_{min} in length. The units are packets.

B.2.5.3 Gap

A gap is the longest sequence of packets beginning and ending with a loss or the start or end of reception, where the sequences of consecutive received packets are greater than or equal to G_{min} in length. The units are packets.

B.2.5.4 Burst density

The burst density is the fraction of RTP data packets within burst periods since the beginning of reception that were either lost or discarded. This value is expressed as a fixed point number with the binary point at the left edge of the field. It is calculated by dividing the total number of packets lost or discarded (excluding duplicate packet discards) within burst periods by the total number of packets expected within the burst periods, multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part. The units are the ratio of lost and discarded packets to total packets.

B.2.5.5 Gap density

The gap density is the fraction of RTP data packets within inter-burst gaps since the beginning of reception that were either lost or discarded. The value is expressed as a fixed point number with the binary point at the left edge of the field. It is calculated by dividing the total number of packets lost or discarded (excluding duplicate packet discards) within gap periods by the total number of packets expected within the gap periods, multiplying the result of the division by 256, limiting the maximum value to 255 (to avoid overflow), and taking the integer part. The units are the ratio of lost and discarded packets to total packets.

B.2.5.6 Burst duration

The burst duration is the mean of the burst periods, expressed in units of milliseconds, which have occurred since the beginning of reception. The duration of each period is calculated based upon the packets that mark the beginning and end of that period. It is equal to the timestamp of the end packet, plus the duration of the end packet, minus the timestamp of the beginning packet. If the actual values are not available, estimated values MUST be used. If there have been no burst periods, the burst duration value MUST be zero.

B.2.5.7 Gap duration

The gap duration is the mean of the gap periods, expressed in units of milliseconds, which have occurred since the beginning of reception. The duration of each period is calculated based upon the packet that marks the end of the prior burst and the packet that marks the beginning of the subsequent burst. It is equal to the timestamp of the subsequent burst packet, minus the timestamp of the prior burst packet, plus the duration of the prior burst packet. If the actual values are not available, estimated values MUST be used. In the case of a gap that occurs at the beginning of reception, the sum of the timestamp of the prior burst packet and the duration of the prior burst packet are replaced by the reception start time. In the case of a gap that occurs at the end of reception, the timestamp of the subsequent burst packet is replaced by the reception end time. If there have been no gap periods, the gap duration value MUST be zero.

B.3 Example packet trace

There are two charts shown below that were obtained from analysis of an example IP trace. The first chart shows a scatter diagram of burst length versus burst weight (Gilbert model). Burst length is the distance in packets between the first and last lost packets in a burst and burst weight is the number of packets lost within the burst. It can be clearly seen that bursts of up to 100 packets in length occur, and have a typical loss density of 20-25%.

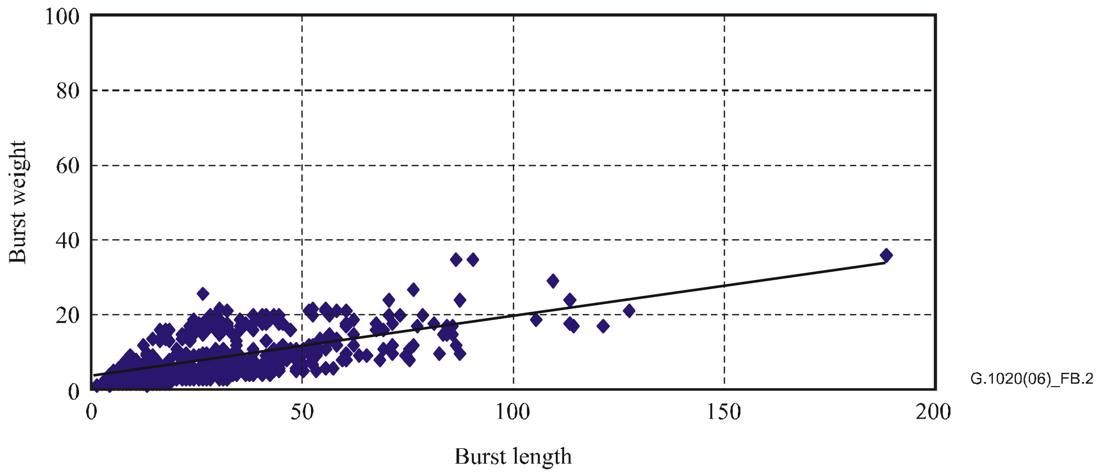

Figure B.2/G.1020 – Trace W3 scatter diagram of burst length vs weight for packet loss only

The second chart shows a scatter diagram of burst length versus burst weight for losses and discards, assuming a 50 ms fixed de-jitter buffer size. This shows a much larger number of bursts indicating that jitter was a significant problem on this trace. Burst density extends out to 500 packets and mean burst density is approximately 30%.

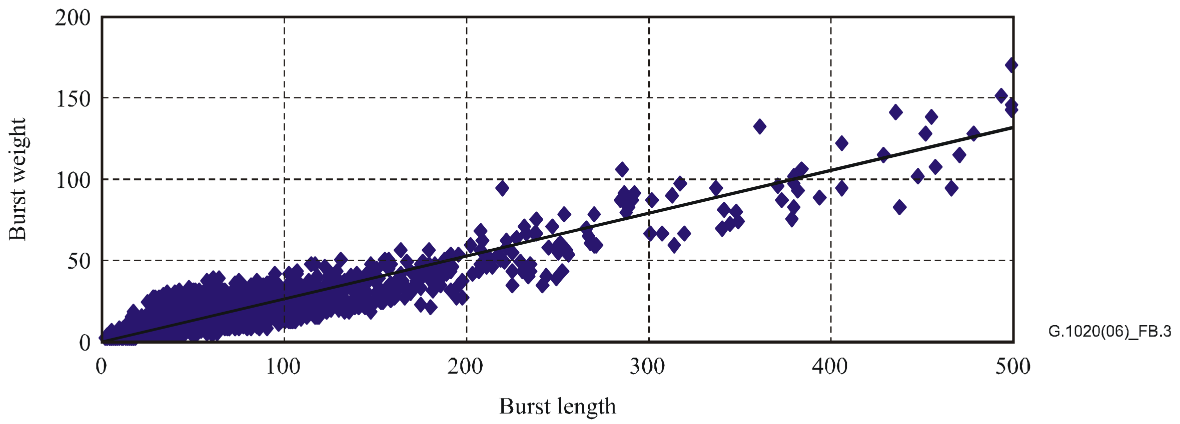

Figure B.3/G.1020 – Trace W3 scatter diagram of burst length vs weight for packet loss and packet discard (50 ms jitter buffer)

B.4 Bibliography to Annex B

[B-1] ALTMAN (E.), AVRACHENKOV (K.), BARAKAT (C.): TCP in the Presence of Bursty Losses, Performance Evaluation 42, pp. 129-147, 2000.

[B-2] BERGER (J.M.), MANDELBROT (B.): A New Model for Error Clustering in Telephone Circuits, IBM J R&D, July 1963.

[B-3] BLANK (H.A.), TRAFTON (P.J.): A Markov Error Channel Model, *Proc. Nat Telecomm Conference*, 1973.

[B-4] BOLOT (J.C.), VEGA GARCIA (A.): The case for FEC based error control for packet audio in the Internet, *ACM Multimedia Systems*, 1997.

[B-5] BOUTREMANS (C.), IANNACCONE (G.), DIOT (C.): Impact of Link Failures on VoIP Performance, *Sprint Labs Technical Report IC/2002/015*.

[B-6] CAIN (J.B.), SIMPSON (R.S.): The Distribution of Burst Lengths on a Gilbert Channel, IEEE Trans IT, 15 September 1969.

[B-7] CLARK (A.): Modelling the effects of burst packet loss and recency on subjective voice quality, *IPtel 2001 workshop*.

[B-8] DRAJIC (D.), VUCETIC (B.): Evaluation of Hybrid Error Control Systems, *IEE Proc. F.* Vol. 131, April 1984.

[B-9] EBERT (J-P.), WILLIG (A.): A Gilbert-Elliott Model and the Efficient Use in Packet Level Simulation, *TKN Technical Report 99-002*.

[B-10] ELLIOTT (E.O.): Estimates of Error Rates for Codes on Burst Noise Channels, *BSTJ 42*, September 1963.

[B-11] ELLIOTT (E.O.): A Model of the Switched Telephone Network for Data Communications, *BSTJ 44*, January 1965.

[B-12] ETSI TIPHON TS 101 329-5 VI.1.2 (2002), *Annex E, QoS measurements for Voice over IP*.

[B-13] GILBERT (E.N.): Capacity of a Burst Noise Channel, *BSTJ*, September 1960.

[B-14] JIANG (W.), SCHULZRINNE (H.): Modelling of Packet Loss and Delay and their effect on Real Time Multimedia Service Quality, *NOSSDAV*, 2000.

[B-15] LEWIS (P.), COX (D.): A Statistical Analysis of Telephone Circuit Error Data, *IEEE Trans COM-14*, 1966.

[B-16] MERTZ (P.): Statistics of Hyperbolic Error Distributions in Data Transmission, *IRE Trans CS-9*, December 1961.

[B-17] SANNECK (H.), CARLE (G.): A Framework Model for Packet Loss Metrics Based on Loss Runlengths, *Proc. ACM MMCN*, January 2000.

[B-18] YAJNIK (M.), MOON (S.), KUROSE (J.), TOWSLEY (D.): Measuring and Modelling of the Temporal Dependence in Packet Loss, *UMASS CMPSCI Tech. Report 98-78*.

Annex C

Example adaptive de-jitter buffer emulator

C.1 Introduction

This example of a de-jitter buffer emulator operates by tracking the short-term minimum delay and using this to position a time window equivalent in size to the de-jitter buffer size. The actual packet arrival time is compared to the time window to determine if the packet would be discarded or accommodated.

The output from this de-jitter buffer emulator is a packet loss/discard event associated with a count of the number of good packets (i.e., not lost or discarded), which is input to the packet loss distribution model.

The de-jitter buffer emulation algorithm determines the delay variation for each arriving RTP packet, based on the RTP timestamp/sequence number and a local clock. This approach is preferable to measuring packet-to-packet delay variation as it:

i) handles out-of-order packets without requiring them to be buffered, which reduces computational complexity;

ii) is able to detect mid-long term delay variations, due to congestion, route changes or timing drift.

C.2 Parameter definitions

C.2.1 (De-)Jitter buffer adaptive

A logical indication (expressed in two bit positions) of whether the de-jitter buffer size is adaptive (11) or fixed/non-adaptive (10). The bit pattern (00) indicates that the mode is unknown, and (01) is reserved.

C.2.2 (De-)Jitter buffer Rate

The buffer adjustment rate is the approximate time it takes for an adaptive buffer to adjust to a step change in delay variation range from 30 ms to 100 ms. The rate, R, is calculated according to the following formula: adjustment time = 2 * R * (packet payload size, in milliseconds).

C.2.3 (De-)Jitter buffer Nominal Delay

This is the current setting of buffering time applied to a packet that arrives exactly on time, meaning that the packet encounters exactly the same transfer delay as the reference packet used to align the de-jitter buffer (the first packet in the emulator below). The unit of measure is milliseconds.

C.2.4 (De-)Jitter buffer Maximum Delay

This is the buffering time applied to a packet that represents the earliest arrival time for a packet that would not be discarded, meaning that the packet encounters a very short transfer delay with respect to the reference packet, but is still buffered and played-out successfully. The unit of measure is milliseconds.

C.2.5 (De-)Jitter buffer Absolute Maximum Delay

This is the largest value of buffering delay that can be applied to a received packet, under adaptive buffer control. With fixed buffer control, this parameter is equal to the Maximum Delay, defined above. This is usually a configuration parameter. The unit of measure is milliseconds.

C.3 De-jitter buffer emulation

The de-jitter buffer emulator operates as follows:

The first arriving RTP packet is the initial reference point with RTP timestamp R_{ref}.

Set Nominal equal to the delay for packets arriving on time (configuration parameter).

Set Maximum Delay equal to the number of packets times the packet size (configuration parameter).

Define early window = Maximum – Nominal.

Define late window = Nominal.

For each RTP packet associated with a stream that passes the monitoring point:

> *Associate a local timestamp L with the arrival time of the RTP packet*
>
> *Identify the RTP timestamp R of the packet*
>
> *Estimate the expected arrival time of the RTP packet based on the reference RTP packet using the expression $L_{expected} = L_{ref} + (R - R_{ref})$*
>
> *Estimate the delay variation of the RTP packet as $D = L - L_{expected}$*
>
> > *If D < early window then*
> >
> > > *mark the packet as discarded*
> > >
> > > *reset the reference point to this packet*
> >
> > *If D > late window then*
> >
> > > *mark the packet as discarded*
>
> *If packet is a duplicate of an already received packet, then silently discard*
>
> *Maintain a sliding window of 32 packets, ordered by sequence number, by default marked as lost – mark packets within this window as accommodated or discarded*
>
> *At the end of the window – identify packets as lost/discarded or accommodated*

The early/late window can be dynamically adjusted to suit adaptive de-jitter buffer behaviour.

Adjustment algorithm:

Define threshold T1 equal to the lowest unacceptable rate of discards (a configurable parameter)

Define threshold T2 equal to the period between de-jitter buffer size downward adjustments (in packets, a configurable parameter)

Maintain a running average of late discards C1, with a scaling factor, S (typically 15)

$C1 = (C1 \times (S - 1) + D)/S$ where D is 1 if packet discarded and 0 if not

Maintain a count of the packets received since the last late discard C2

if C1 exceeds a threshold, T1, and the buffer is less than the maximum then increase the buffer size, and reset C1;

if C2 exceeds a threshold, T2, and the buffer is more than the minimum then reduce the buffer size, and reset C2.

The maximum value of the time window, or de-jitter buffer maximum length, must be specified so that the emulator cannot grow the buffer to extreme values that would not be possible in practice.

Appendix I

List of RTCP XR metrics

This appendix gives the mapping between RTCP-XR metrics and various ITU-T Recommendations, with a detailed description of the action taken to include these metrics in the main body of the Recommendation.

RTCP XR Metrics	Existing ITU-T references	Action
Loss rate	Y.1540: IPLR is at the IP layer, need new definition for the RTP layer computed at receiver-only.	Definition added to 7.2.2, based on 4.7.1/RFC 3611. There are 2 alternative frameworks included for other definitions needed, such as the definition of a lost packet.
Discard rate	Mentioned in G.1020 Overall Loss, but not defined.	Definition added to 7.2.2, based on 4.7.1/RFC 3611.
burst density	The model is already defined in Annex B.	Define a loss model, based on the Gilbert-Elliott model. Definitions in B.2.5.
gap density		
burst duration		
gap duration		
G_{min}		
Round trip delay		Take the definition of RFCs 3550 and 3611 (Needs refinement, added to 8.3)
End-system delay	Source Terminal and Destination Terminal Delay of G.1020 sum to this value.	Need to work on a precise definition (based on the one in RFC 3611, added to 8.2).
Signal level	P.56, P.561	Promote the use of these standard methods.
Noise level	O.41, P.561	
Residual echo return loss	G.122 for echo loss, G.168 (SG 16) for residual echo level	
R factor	G.107	Need for specific default values
External R factor	G.107	Need for specific default value
MOS LQ	P.862, P.563, P.564	A field is missing in RTCP XR VoIP Metrics Report Block to define what model has been used
MOS CQ	P.562 (CCI), G.107	A field is missing in RTCP XR VoIP Metrics Report Block to define what model has been used
Packet loss concealment	G.1020 has material on this in 7.4.	
Jitter buffer adaptive	Clause 7.2 covers most of the jitter buffer parameters.	
Jitter buffer rate		There is a de-jitter buffer model in Annex C, and definitions have been added. (This was formerly App. II)
Jitter buffer nominal delay		
Jitter buffer maximum delay		
Jitter buffer absolute maximum delay		

BIBLIOGRAPHY

[1] [RFC 3550], RTP (2003), *A Transport Protocol for Real-Time Applications*.

[2] [RFC 2330] (1998), *Framework for IP Performance Metrics*.

[3] [RFC 2678] (1999), *IPPM Metrics for Measuring Connectivity*.

[4] [RFC 2679] (1999), *A One-way Delay Metric for IPPM*.

[5] [RFC 2680] (1999), *A One-way Packet Loss Metric for IPPM*.

[6] [RFC 2681] (1999), *A Round-trip Delay Metric for IPPM*.

[7] [RFC 3357] (2002), *One-way Loss Pattern Sample Metrics*.

[8] [RFC 3393] (2002), *IP Packet Delay Variation Metric for IP Performance Metrics IPPM*.

[9] [RFC 3432] (2002), *Network performance measurement for periodic streams*.

[10] ETSI [TS 101 329-5] V1.1.2 (2002), *TIPHON release 3, End-to-End Quality of Service in TIPHON Systems, Part 5 Quality of Service (QoS) Measurement Methodologies*.

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, 2007