INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.114

(05/2003)

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

International telephone connections and circuits – General Recommendations on the transmission quality for an entire international telephone connection

One-way transmission time

ITU-T Recommendation G.114

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
General definitions	G.100–G.109
General Recommendations on the transmission quality for an entire international telephone connection	**G.110–G.119**
General characteristics of national systems forming part of international connections	G.120–G.129
General characteristics of the 4-wire chain formed by the international circuits and national extension circuits	G.130–G.139
General characteristics of the 4-wire chain of international circuits; international transit	G.140–G.149
General characteristics of international telephone circuits and national extension circuits	G.150–G.159
Apparatus associated with long-distance telephone circuits	G.160–G.169
Transmission plan aspects of special circuits and connections using the international telephone connection network	G.170–G.179
Protection and restoration of transmission systems	G.180–G.189
Software tools for transmission systems	G.190–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
TESTING EQUIPMENTS	G.500–G.599
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
DIGITAL TERMINAL EQUIPMENTS	G.7000–G.7999
DIGITAL NETWORKS	G.8000–G.8999

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

ITU-T Recommendation G.114

One-way transmission time

Summary

This Recommendation provides guidance on the effect of end-to-end one-way delay (sometimes termed latency), and an upper bound one-way network delay.

While it is recommended that a one-way delay of 400 ms should not be exceeded for general network planning, it is important to appreciate that highly interactive tasks (e.g., many voice calls, interactive data applications, video conferencing) can be affected by much lower delays.

The effects of delays below 500 ms on conversational speech are estimated using a curve derived from the E-model (ITU-T Rec. G.107).

This version constitutes a major revision of this Recommendation in order to align with other ITU-T Recommendations of the G.100 series.

Source

ITU-T Recommendation G.114 was approved by ITU-T Study Group 12 (2001-2004) under the ITU-T Recommendation A.8 procedure on 6 May 2003. It includes the modifications introduced by ITU-T Rec. G.114 (2003) Appendix II approved on 30 September 2003.

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, 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 G.114

One-way transmission time

1 Introduction

This Recommendation provides guidance on the effect of end-to-end one-way delay (sometimes termed latency), and an upper bound on one-way network delay. The effect of delay on speech transmission quality can be estimated by the use of a curve derived from the Transmission Rating Model of ITU-T Rec. G.107 [3], which is the recommended ITU-T method for end-to-end speech transmission planning. ITU-T Rec. G.108 [4] gives detailed examples on how to use the model to assess the transmission performance of connections involving various impairments, including one-way delay; and ITU-T Rec. G.109 [5] maps transmission rating predictions of the model into categories of speech transmission quality. Thus, while ITU-T Rec. G.114 provides useful information regarding one-way delay as a parameter by itself, ITU-T Rec. G.107 [3] (and its ITU-T Rec. G.108 [4] and ITU-T Rec. G.109 [5] companions) should be used to assess the effects of delay in conjunction with other impairments (e.g., distortions due to speech processing).

Highly interactive tasks (e.g., some speech, video conferencing and interactive data applications) may be affected by delays below 100 ms, as per test result documented in Annex B of previous versions of ITU-T Rec. G.114. For this reason, previous versions of this Recommendation noted that if delays were kept below 150 ms, then most *applications* would not be significantly affected. Additionally, an upper limit of 400 ms for *network* planning purposes was always a part of ITU-T Rec. G.114. However, this parallel treatment of network delays on one hand, with application ("mouth-to-ear") level delays on the other hand, led to confusion in how ITU-T Rec. G.114 should be applied.

Fortunately, with the development and approval of the E-model (ITU-T Rec. G.107 [3]), which is based on subjective tests of delay (among other parameters), there now exists an agreed way of estimating the effects of delay on mouth-to-ear speech transmission quality.

Accordingly, simple and straightforward guidance can now be provided for the effects of delay on speech transmission and is given in the present Recommendation.

The lack of similar tools for non-speech applications is a subject for further study, so this Recommendation can only provide general planning guidance.

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.

[1] ITU-T Recommendation G.100 (2001), *Definitions used in Recommendations on general characteristics of international telephone connections and circuits*.

[2] ITU-T Recommendation G.101 (1996), *The transmission plan*.

[3] ITU-T Recommendation G.107 (2003), *The E-Model, a computational model for use in transmission planning*.

[4] ITU-T Recommendation G.108 (1999), *Application of the E-model: A planning guide*.

[5] ITU-T Recommendation G.109 (1999), *Definition of categories of speech transmission quality*.

[6] ITU-T Recommendation G.131 (1996), *Control of talker echo*.

[7] ITU-T Recommendation G.168 (2002), *Digital network echo cancellers*.

[8] ITU-T Recommendation G.763 (1998), *Digital circuit multiplication equipment using G.726 ADPCM and digital speech interpolation*.

[9] ITU-T Recommendation G.764 (1990), *Voice packetization – Packetized voice protocols*.

[10] ITU-T Recommendation G.766 (1996), *Facsimile demodulation/remodulation for digital circuit multiplication equipment*.

[11] ITU-T Recommendation G.767 (1998), *Digital circuit multiplication equipment using 16 kbit/s LD-CELP, digital speech interpolation and facsimile demodulation/remodulation*.

[12] ITU-T Recommendation Q.551 (2002), *Transmission characteristics of digital exchanges*.

[13] ITU-T Recommendation Y.1541 (2002), *Network performance objectives for IP-based services*.

3 Applicability to speech transmission quality – Use of the E-model

This Recommendation provides end-to-end limits for mean one-way delay, independent of other transmission impairments. The need to consider the combined effects of all impairments on speech transmission quality is addressed by the Transmission Rating Model of ITU-T Rec. G.107 [3], which is the recommended ITU-T method for end-to-end speech transmission planning. ITU-T Rec. G.108 [4] gives detailed examples on how to use the model to assess the transmission performance of connections involving various impairments, including delay; and ITU-T Rec. G.109 [5] maps transmission rating predictions of the model into categories of speech transmission quality. Thus, while this Recommendation provides useful information regarding mean one-way delay as a parameter by itself, ITU-T Rec. G.107 [3] (and its ITU-T Rec. G.108 [4] and ITU-T Rec. G.109 [5] companions) should be used to assess the effects of delay in conjunction with other impairments (e.g., distortions due to speech processing).

4 Recommendations for one-way transmission time

Regardless of the type of application, it is recommended to not exceed a one-way delay of 400 ms for general network planning (i.e., UNI to UNI, as illustrated, for example, in ITU-T Rec. Y.1541 [13]), a value that allows flexibility in deploying global networks, without making an excessive number of user experiences unacceptable.

However, it is desirable to keep the delays seen by user applications as low as possible. The E-model should be used to estimate the effect of one-way delay (including all delay sources, i.e., "mouth-to-ear") on speech transmission quality for conversational speech as shown below. For non-speech applications such as interactive data or video, there are no agreed-upon assessment tools like the E-model, so the effects of delay on such applications must be carefully monitored. Although a few applications may be slightly affected by end-to-end (i.e., "mouth-to-ear" in the case of speech) delays of less than 150 ms, if delays can be kept below this figure, most applications, both speech and non-speech, will experience essentially transparent interactivity.

While delays above 400 ms are unacceptable for general network planning purposes, it is recognized that in some exceptional cases this limit will be exceeded. An example of such an exception is an unavoidable double satellite hop for a hard-to-reach location, the impact of which can be estimated by use of the advantage factor in the E-model.

Regarding the use of the E-model for speech applications, the effect of delay can be seen in the following graph of Transmission Rating, R, versus delay. Also shown are the speech quality categories of ITU-T Rec. G.109 [5], which translate the R values to levels of user acceptance.

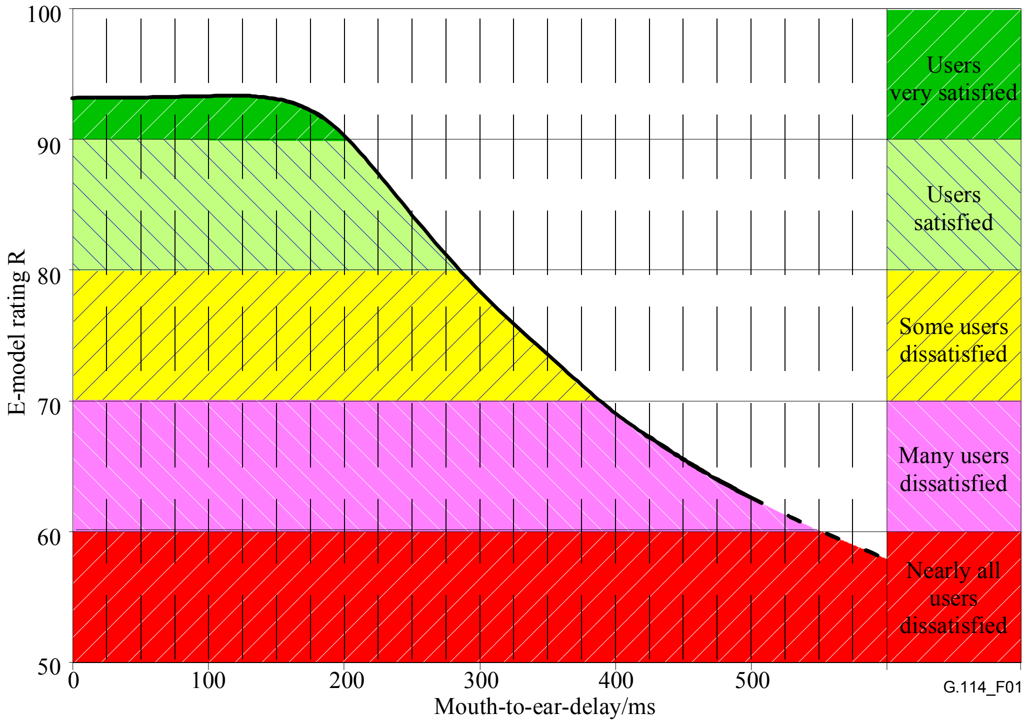

Figure 1/G.114 – Determination of the effects of absolute delay by the E-model

NOTE 1 – The curve in Figure 1 is based on the effect of pure delay only, i.e., in the complete absence of any echo. This is calculated by setting the G.107 E-model parameter Ta equal to the total value of one-way delay from mouth-to-ear, with all other E-model input parameter values set to their default values. The effect of echo, as would be incurred due to imperfect echo control, will result in lower speech quality for a given value of one-way delay.

NOTE 2 – The calculation also assumes an Equipment Impairment Factor (Ie) of zero. Non-zero values, as would be incurred due to speech coding/processing, will result in lower speech quality for a given value of one-way delay.

NOTE 3 – For one-way delay values exceeding 500 ms, the graph is continued as a dashed line to indicate that these results are not fully validated, but is the best estimate of what should be expected, and, therefore, provides useful guidance.

5 Estimating end-to-end delay based on assemblies of transmission elements

The nominal delay values and general planning rules given in Annex A, and the coder-related delays of Appendix I, may be used to estimate the total end-to-end transmission time.

Annex A

End-to-end delay estimation

A.1 Planning values for the delay of transmission elements

Table A.1/G.114 – Planning values for the delay of transmission elements

Transmission or processing system	Contribution to one-way transmission time	Remarks
Terrestrial coaxial cable or radio-relay system: FDM and digital transmission	4 µs/km	Allows for delay in repeaters and regenerators
Optical fibre cable system, digital transmission	5 µs/km (Note 1)	
Submarine coaxial cable system	6 µs/km	
Submarine optical fibre system: – transmit terminal – receive terminal	 13 ms 10 ms	Worst case
Satellite system: – 400 km altitude – 14 000 km altitude – 36 000 km altitude	 12 ms 110 ms 260 ms	Propagation through space only (between earth stations)
FDM channel modulator or demodulator	0.75 ms (Note 2)	
PLMS (Public Land Mobile System) – objective 40 ms	80-110 ms	
H.260-series video coders and decoders	Further study (Note 3)	
DCME (ITU-T Rec. G.763 [8]) per pair: for speech, VBD, and non-remodulated fax	30 ms	Half the sum of transmission times in both directions of transmission
DCME (ITU-T Rec. G.767 [11]) per pair: for speech, VBD, and non-remodulated fax	30 ms	
DCME (ITU-T Rec. G.766 [10] in conjunction with ITU-T Recs G.763 [8] or G.767 [11]) per pair: for remodulated fax	200 ms	
PCME (ITU-T Rec. G.764 [9]) per pair: – with speech and non-remodulated VBD – with remodulated VBD	 35 ms 70 ms	
Transmultiplexer	1.5 ms (Note 4)	Half the sum of transmission times in both directions of transmission
Digital transit exchange, digital-digital	0.45 ms (Note 5)	
Digital local exchange, analogue-analogue	1.5 ms (Note 5)	
Digital local exchange, analogue subscriber line-digital junction	0.975 ms (Note 5)	
Digital local exchange, digital subscriber line-digital junction	0.825 ms (Note 5)	
Echo cancellers	0.5 ms (Note 6)	
ATM (CBR using AAL 1)	6.0 ms (Note 7)	

Table A.1/G.114 – Planning values for the delay of transmission elements

NOTE 1 – This value is provisional and is under study.
NOTE 2 – These values allow for group-delay distortion around frequencies of peak speech energy and for delay of intermediate higher order multiplex and through-connecting equipment.
NOTE 3 – Further study required. Delay for these devices is usually non-constant, and the range varies by implementation. Current implementations are of the order of several hundred milliseconds, and considerable delay is added to audio channels to achieve lip-synchronization. Manufacturers are encouraged to reduce their contribution to transmission time, in accordance with this Recommendation.
NOTE 4 – For satellite digital communications where the transmultiplexer is located at the earth station, this value may be increased to 3.3 ms.
NOTE 5 – These are mean values: depending on traffic loading, higher values can be encountered, e.g., 0.75 ms (1.950 ms, 1.350 ms, or 1.250 ms) with 0.95 probability of not exceeding. (For details see ITU-T Rec. Q.551 [12].)
NOTE 6 – This is averaged for both directions of transmission.
NOTE 7 – This is the cell formation delay of 64 kbit/s stream when it completely fills the cell (one voice channel per VC). In practical applications, additional delay will result, e.g., from cell loss detection and buffering. Other delays may be applicable to other AALs and cell mapping arrangements, and are for further study.

A.2 Codec delay

Modern speech codecs operate on collections of speech samples known as frames. Each block of input speech samples is processed into a compressed frame. The coded speech frame is not generated until all speech samples in the input block have been collected by the encoder. Thus, there is a delay of one frame before processing can begin. In addition, many coders also look into the succeeding frame to improve compression efficiency. The length of this advance look is known as the look-ahead time of the coder. The time required to process an input frame is assumed to be the same as the frame length since efficient use of processor resources will be accomplished when an encoder/decoder pair (or multiple encoder/decoder pairs operating in parallel on multiple input streams) fully uses the available processing power (evenly distributed in the time domain). Thus, the delay through an encoder/decoder pair is normally assumed to be:

$$2 \times \text{frame size} + \text{look-ahead}$$

A.2.1 Delay in wirebound environment

If the output facility is running at the same rate as the speech codec (e.g., an 8 kbit/s facility for ITU-T Rec. G.729), then an additional frame of delay is incurred when clocking the compressed frame to the facility. Thus, the maximum delay attributable to codec-related processing in conventional wirebound systems (i.e., the PSTN) is:

$$3 \times \text{frame size} + \text{look-ahead}$$

A.2.2 Delay in mobile and wireless environment

If the output facility is a mobile network or a cordless facility, then the frame output by the encoder will function similar to the operation in wirebound environment but an additional delay is incurred for attaching the compressed frame to the airpath (assumed again that the mobile facility is running at the same rate as the speech codec). Thus, the maximum delay attributable to codec-related processing in mobile and wireless systems is:

$$3 \times \text{frame size} + \text{look-ahead} + \text{air interface framing}$$

A.2.3 Delay in IP environment (one frame per packet)

If the output facility is an IP network, then the frame output by the encoder will instantaneously be dropped into an IP packet. The additional delay required for IP packet assembly and presentation to the underlying link layer will depend on the link layer. When the link layer is a LAN (e.g., Ethernet), this additional time will usually be quite small. Thus, the minimum delay attributable to codec-related processing in IP-based systems is:

$$2 \times \text{frame size} + \text{look-ahead}$$

When the link layer is one with lower clock rate (e.g., Modem connection) or one with high traffic load (e.g., congested LAN), the additional delay will increase substantially. In order to clock compressed frames at least with the same rate to the facility as the speech samples are collected at the input of the encoder, the additional delay should not exceed one frame size. Thus, the maximum delay attributable to codec-related processing in IP-based systems operating in real-time is:

$$3 \times \text{frame size} + \text{look-ahead}$$

A.2.4 Delay in IP environment (multiple frames per packet)

If multiple voice frames are grouped together into a single IP packet, further delay is added to the speech signal. This delay will be at least the duration of one extra voice frame at the encoder for each additional voice frame added to the IP packet. Thus, the minimum delay attributable to codec-related processing in IP-based systems with multiple frames per packet is:

$$(N + 1) \times \text{frame size} + \text{look-ahead}$$

where N is the number of frames in each packet.

When the link layer is one with lower clock rate (e.g., Modem connection) or one with high traffic load (e.g., congested LAN), additional delay will be incurred in delivering the packet to the facility. In order to clock compressed frames at least with the same rate to the facility as the speech samples are collected at the input of the encoder, the additional delay should, in case of multiple frames per packet, not exceed the length of the frames contained in one packet. It should be noted that clocking out a packet to the IP facility cannot start before all speech frames for this packet are available. Thus, the maximum delay attributable to codec-related processing in IP-based systems operating in real-time with multiple frames per packet is:

$$(2N + 1) \times \text{frame size} + \text{look-ahead}$$

where N is the number of frames in each packet.

The Figure A.1 provides an example for N = 2:

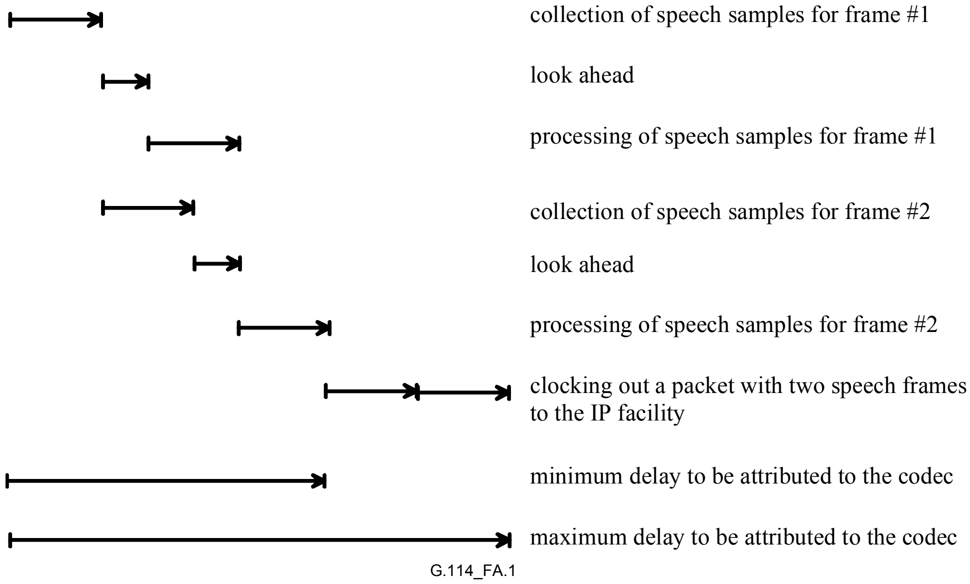

Figure A.1/G.114 – Example: Composition of total codec-related delay in an IP Environment for N = 2

A.3 Delay due to IP delay variation buffer

Packetized transmission systems exhibit variable delay (jitter) in packet delivery time. This is caused by the fact that different packets carrying speech samples of the same telephone conversation may encounter varying queue lengths, or different routes through the network. Details of this effect depend strongly on the specific mechanisms for transport, queuing or prioritization, which may be implemented in such a system. Nevertheless, delay variation must be removed prior to replaying the speech to the user, otherwise significant degradation will be noticeable.

This is typically achieved by collecting packets in a de-jitter buffer at the receive side. This buffer rearranges the timely order of the packets, and is sized to take into account a certain range of network delay variation, effectively delaying all packets to correspond to the delay of the packet with the longest tolerable transit time. If the delivery time of a packet exceeds the length of the receive buffer, then this packet "arrives too late" with respect to its intended play-out time, and will be discarded. Consequently, the speech carried in this packet is lost for the decoding process. This "packet loss" impairs speech transmission quality (see ITU-T Rec. G.113).

The contribution of the de-jiffer buffer to one-way delay is based on the average time packets spend in the buffer, which is less than the peak buffer size. Depending on the specific type of implementation, as well as on the proper adjustment of the de-jitter buffer, this may be as low as one half of the peak buffer size (assuming symmetrically distributed delays). Packets that encounter the minimum transfer delay will wait the maximum time in the de-jitter buffer before being played out as a synchronous stream, while the reverse is true for packets with the maximum accommodated transfer delay (these packets spend the minimum time in the de-jitter buffer). For PLANNING PURPOSES, IT IS RECOMMENDED to assume that, a de-jitter buffer adds one half of its peak delay to the mean network delay.

Example (taken from Appendix III/Y.1541 [13]):

A jitter buffer designed to compensate for 50 ms packet delay variation range will introduce 25 ms additional delay, on average.

Further Guidance on the effects of packet delay, as caused by de-jitter buffer, is provided by ITU-T Rec. Y.1541 [13].

It should be noted that, with the use of dynamic de-jitter buffer implementations, the delay of the speech replayed to the user will be subjected to infrequent transitional delay variations when the de-jitter buffer resizes.

Appendix I

Delay introduced by coder-related processing

Table I.1/G.114 – Delay values for coders in wirebound applications

Coder type	Rate (kbit/s)	Frame size (ms)	Look-ahead (ms)	Mean one-way delay introduced by coder-related processing (ms)	Reference
PCM	64	0.125	0	0.375	G.711, G.712
ADPCM	40	0.125	0	0.375	G.726, G.727
ADPCM	32	0.125	0	0.375	G.721(1988), G.726, G.727
ADPCM	24	0.125	0	0.375	G.726, G.727
ADPCM	16	0.125	0	0.375	G.726, G.727
LD-CELP	16	0.625	0	1.875	G.728
LD-CELP	12.8	0.625	0	1.875	G.728
CS-ACELP	8	10	5	35	G.729
VSELP	7.95	20	0	60	IS-54-B, TIA
ACELP	7.4	20	5	65	IS-641, TIA
QCELP	8	20	0	60	IS-96-A
RCELP	8	20	10	70	IS-127
VSELP	6.7	20	5	65	Japanese PDC
RPE-LTP	13	20	0	60	GSM 06.10, Full-rate
VSELP	5.6	20	0	60	GSM 06.20, Half-rate
ACELP	12.2	20	0	60	GSM 06.60, Enhanced FR
ACELP	5.3	30	7.5	97.5	G.723.1
MP-MLQ	6.3	30	7.5	97.5	G.723.1

NOTE 1 – The PCM coder converts from analogue to digital and vice-versa while all other coders refer to the PCM domain; for PCM in the analogue domain, additional delay is incurred (0.375 ms).

NOTE 2 – For wirebound applications, the mean one-way delay introduced by codec-related processing = 3 × frame size + look-ahead (see A.2.1).

Table I.2/G.114 – Delay values for coders in mobile or cordless applications

Coder type	Rate (kbit/s)	Frame size (ms)	Look-ahead (ms)	Air interface framing (ms)	Mean one-way delay introduced by coder-related processing (ms)	Reference
PCM	64	0.125	0	(See Note 3)		G.711, G.712
ADPCM	40	0.125	0	(See Note 3)		G.726, G.727
ADPCM	32	0.125	0	13.625	14	G.721(1988), G.726, G.727, DECT
ADPCM	24	0.125	0	(See Note 3)		G.726, G.727
ADPCM	16	0.125	0	(See Note 3)		G.726, G.727
LD-CELP	16	0.625	0	(See Note 3)		G.728
LD-CELP	12.8	0.625	0	(See Note 3)		G.728
CS-ACELP	8	10	5	(See Note 3)		G.729
VSELP	7.95	20	0			IS-54-B, TIA
ACELP	7.4	20	5			IS-641, TIA
QCELP	8	20	0			IS-96-A
RCELP	8	20	10			IS-127
VSELP	6.7	20	5			Japanese PDC
RPE-LTP	13	20	0	35	95	GSM 06.10, Full-rate
VSELP	5.6	20	0	35	95	GSM 06.20, Half-rate
ACELP	12.2	20	0	35	95	GSM 06.60, Enhanced FR
ACELP	5.3	30	7.5	(See Note 3)		G.723.1
MP-MLQ	6.3	30	7.5	(See Note 3)		G.723.1

NOTE 1 – The PCM coder converts from analogue to digital and vice-versa while all other coders refer to the PCM domain; for PCM in the analogue domain, additional delay is incurred (0.375 ms).

NOTE 2 – For mobile or cordless applications the mean one-way delay introduced by codec-related processing = 3 × frame size + look-ahead + air interface framing (see A.2.2)

NOTE 3 – For the marked types of coders, Study Group 12 is not aware of any mobile or cordless application.

Table I.3/G.114 – Delay values for coders in IP-based applications (one frame per packet)

Coder type	Rate (kbit/s)	Frame size (ms)	Look-ahead (ms)	Mean one-way delay introduced by coder-related processing (ms) (see Note 2)		Reference
				Minimum	Maximum	
PCM	64	0.125	0	0.25	0.375	G.711, G.712
ADPCM	40	0.125	0	0.25	0.375	G.726, G.727
ADPCM	32	0.125	0	0.25	0.375	G.721(1988), G.726, G.727
ADPCM	24	0.125	0	0.25	0.375	G.726, G.727
ADPCM	16	0.125	0	0.25	0.375	G.726, G.727
LD-CELP	16	0.625	0	1.25	1.875	G.728
LD-CELP	12.8	0.625	0	1.25	1.875	G.728
CS-ACELP	8	10	5	25	35	G.729
VSELP	7.95	20	0	40	60	IS-54-B, TIA
ACELP	7.4	20	5	45	65	IS-641, TIA
QCELP	8	20	0	40	60	IS-96-A
RCELP	8	20	10	50	70	IS-127
VSELP	6.7	20	5	45	65	Japanese PDC
RPE-LTP	13	20	0	40	60	GSM 06.10, Full-rate
VSELP	5.6	20	0	40	60	GSM 06.20, Half-rate
ACELP	12.2	20	0	40	60	GSM 06.60, Enhanced FR
ACELP	5.3	30	7.5	67.5	97.5	G.723.1
MP-MLQ	6.3	30	7.5	67.5	97.5	G.723.1

NOTE 1 – The PCM codec converts from analogue to digital and vice-versa while all other coders refer to the PCM domain; for PCM in the analogue domain, additional delay is incurred (0.375 ms).

NOTE 2 – For IP-related applications, the mean one-way delay introduced by codec-related processing:
= 2 × frame size + look-ahead (minimum, see A.2.3)
= 3 × frame size + look-ahead (maximum, see A.2.3).

Table I.4/G.114 – Delay values for coders in IP-based applications (multiple frames per packet)

Coder type	Rate (kbit/s)	Frame size (ms)	Look-ahead (ms)	Mean one-way delay introduced by coder-related processing (ms) (see Note 2)		Reference
				Minimum	Maximum	
PCM	64	0.125	0	(N + 1) × 0.125	(2N + 1) × 0.125	G.711, G.712
ADPCM	40	0.125	0	(N + 1) × 0.125	(2N + 1) × 0.125	G.726, G.727
ADPCM	32	0.125	0	(N + 1) × 0.125	(2N + 1) × 0.125	G.721(1988), G.726, G.727
ADPCM	24	0.125	0	(N + 1) × 0.125	(2N + 1) × 0.125	G.726, G.727
ADPCM	16	0.125	0	(N + 1) × 0.125	(2N + 1) × 0.125	G.726, G.727
LD-CELP	16	0.625	0	(N + 1) × 0.625	(2N + 1) × 0.625	G.728
LD-CELP	12.8	0.625	0	(N + 1) × 0.625	(2N + 1) × 0.625	G.728
CS-ACELP	8	10	5	(N + 1) × 10 + 5	(2N + 1) × 10 + 5	G.729
VSELP	7.95	20	0	(N + 1) × 20	(2N + 1) × 20	IS-54-B, TIA
ACELP	7.4	20	5	(N + 1) × 20 + 5	(2N + 1) × 20 + 5	IS-641, TIA
QCELP	8	20	0	(N + 1) × 20	(2N + 1) × 20	IS-96-A
RCELP	8	20	10	(N + 1) × 20 + 10	(2N + 1) × 20 + 10	IS-127
VSELP	6,7	20	5	(N + 1) × 20 + 5	(2N + 1) × 20 + 5	Japanese PDC
RPE-LTP	13	20	0	(N + 1) × 20	(2N + 1) × 20	GSM 06.10, Full-rate
VSELP	5.6	20	0	(N + 1) × 20	(2N+1) x 20	GSM 06.20, Half-rate
ACELP	12.2	20	0	(N + 1) × 20	(2N + 1) × 20	GSM 06.60, Enhanced FR
ACELP	5.3	30	7.5	(N + 1) × 30 + 7.5	(2N + 1) × 30 + 7.5	G.723.1
MP-MLQ	6.3	30	7.5	(N + 1) × 30 + 7.5	(2N + 1) × 30 + 7.5	G.723.1

NOTE 1 – The PCM codec converts from analogue to digital and vice-versa while all other coders refer to the PCM domain; for PCM in the analogue domain, additional delay is incurred (0.375 ms).

NOTE 2 – For IP-related applications with multiple frames per packet, the mean one-way delay introduced by codec-related processing can be calculated as follows:
= (N + 1) × frame size + look-ahead (minimum, see A.2.4)
= (2N + 1) × frame size + look-ahead (maximum, see A.2.4).

NOTE 3 – N = number of frames per packet.

Appendix II

Guidance on one-way delay for voice over IP

II.1 Introduction

This appendix gives additional guidance on the application of ITU-T Rec. G.114. The main purpose is to provide practical information for end-to-end VoIP network planning. Also, this appendix provides a linkage to the IP network delay objectives in ITU-T Rec. Y.1541.

II.2 Achieving satisfactory delay

For many *intra-regional* (e.g., within Africa, Europe, North America) routes in the range of 5000 km or less, users of VoIP connections are likely to experience mouth-to-ear delays <150 ms. Appendix III/Y.1541 illustrates this calculation using reference terminals with a total of 50 ms mean delay (10 ms packets). The calculation shows that the 100 ms objective of Y.1541 Class 0 can be met with a well-engineered access network (with a T1 or E1 rate or larger as Y.1541 requires) and with as many as 12 network routers. Appendix X/Y.1541 shows that similar speech quality can be maintained with reference terminals contributing a total delay of a less stringent 80 ms (using 20 ms packets and robust packet loss concealment).

For *inter-regional* routes covered terrestrially, even those traversing the 27 500 km of the ITU's traditional worst-case Hypothetical Reference Connection, a VoIP mouth-to-ear path is likely to see a delay of just over 300 ms. This assumes terminals contributing a total of 80 ms delay (20 ms packets), a well-engineered access network and supporting IP network paths encountering 20 or fewer network routers (as per Appendix III/Y.1541). Of course, it is extremely unlikely that the worst case of 27 500 km will be encountered by many calls. For the much more frequent inter-regional calls of, for example, 10 000 km or less, the corresponding delays would be approximately 225 ms; certainly not as low (or as desirable) as 150 ms, but still quite satisfactory for the vast majority of users.

Whilst delays in the mid-200 ms range may not be a serious problem for long inter-regional calls, where users expect calls to be somewhat different from regional calls, it is critical that network planners do not allow local and regional calls to encounter such delays because user expectations are that such calls be completely delay-transparent.

Whilst it is recognized that using VoIP technologies will increase the delays well above that of non-packetized TDM transmission, this analysis demonstrates that the widespread use of end-end VoIP need not cause problematic delays if appropriate care and planning is exercised.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series B	Means of expression: definitions, symbols, classification
Series C	General telecommunication statistics
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	TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
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 and open system communications
Series Y	Global information infrastructure and Internet protocol aspects
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2003