ANNEX 32

(to the Report of SG16)

QUESTION:	ITU-T Q20/16; ETSI SMG11
SOURCE:	ITU-T SG 16 (Geneva, 7-18 February 2000) (Q.20/16)
TITLE:	Communication on 16 kbit/s wideband speech coding

COMMNUNICATION

FROM:	ITU-T Q. 21/16
TO:	ETSI SMG11
APPROVAL:	Approved by SG 16 (Geneva, February 7-18 2000)
FOR:	Action
DEADLINE:	November 2000

Q20/16:
Mr. R. Drogo De Iacovo, CSELT
Via G. Reiss Romoli, 274
10148 Torino
ITALY
Tel: +39 11 228 6221
Fax: +39 11 228 6207
e-mail: rosario.drogodeiacovo@cselt.it

ETSI/SMG11: Kari Jarvinen
Nokia Research Center
P.O.Box 100 (Sinitaival 6)
FIN-33721 Tampere
FINLAND
Tel: +358-3272-5854
Fax: +358-3272-5888
E-mail:kari.ju.jarvinen@nokia.com

ITU-T Study Group 16 (*Multimedia Services and Systems*) would like to thank ETSI SMG11 for their communication about collaboration on wideband speech codec development. ITU-T SG16 Q.20/16 would like to continue the exchange of information. During our recent meeting in Geneva, February 7-18, 2000, Q.20/16 overall feeling was that it is very important to harmonize the activities between the two bodies in this area.

Q.20/16 group reviewed the ITU-T Terms of Reference for wideband coding around 16 kbit/s with the aim to align them, as far as possible, with the provided version of the AMR wideband performance requirements and design constraints. At the same time Q.20/16 had the opportunity to notice some discrepancies between the ITU-T and the 3GPP-ETSI SMG11 set of requirements and would like to kindly ask 3GPP-ETSI SMG11 to possibly increase the alignment as regards the specific topics reported in the following. Q.20/16 experts group believes that any action taken to increase the commonalities of the two activities, in terms of performance requirements, design constraints and test methodology, is in the direction of potentially leading to harmonized solutions. The ITU-T ToR are annexed to this document.

The areas where discrepancies are found between the ITU-T and the 3GPP-ETSI performance requirements are:

- Talker dependency
- Transmission of DTMF
- Types of background noise. ITU-T is asking SQEG/SG12 for guidance about the appropriate S/N values and measurement methods.

Another area where description of requirements differs (but are not necessarily conflicting) is on the error performance, where 3GPP-ETSI uses a combined source and channel coding, while Q.20/16 considers only the source coding. This discrepancy could be resolved by producing a mapping of the residual error for C/I conditions into our equivalent BER/FER figures.

In the revised ToR, the requirement for bit-rates of 16 and 24 kbit/s were revised to read “not more than” thus allowing for solutions that do not operate at exactly those bit-rates.

Q.20/16 will continue with its workplan, aiming at review of Qualification Test results in its November 2000 meeting. At that date, Q.20/16 will consider the selected AMR-WB algorithm as a participant in the ITU-T selection test, provided that supporting test results (e.g. the AMR-WB selection test results) be made available to Q.20/16. This conclusion has been drawn by comparing the two Terms of Reference and noting that it is likely that the 3GPP-ETSI selected candidate meeting all the 3GPP-ETSI requirements will fulfill, at the same time, the ITU-T requirements, excluding a few areas, as detailed before. It would be useful if a measurement of the residual errors for the C/I conditions be also provided. Unfortunately, as no ITU-based pool funding is available, cost-sharing with other participants is expected.

Please note that most participants in the selection test are expected to fully exploit the available bit-rate, hence we would strongly encourage that the AMR solution also support the 16 and 24 kbit/s rates (although this is not a requirement, as mentioned before). This would maximize the chances to arrive at a harmonized wideband solution.

ITU-T is pleased to continue that collaboration and will consider any further suggestion for the purpose of reaching the desired harmonization.

Attachment:

- ToR for wideband coding around 16 kbit/s (TD (P) 66 Annex Q20.A).

Annex Q20.A
Terms of Reference for ITU-T Wideband (7 kHz) Speech Coding around 16 kbit/s

- Background

The following general guidelines are considered relevant for this wideband activity:

- Input and output audio signals should have a bandwidth of 7 kHz at a sampling rate of 16 kHz.
- Primary signals of interest are clean speech and speech in background noise. Music performance objectives set at higher bit-rates (24 kbit/s).
- High speech quality with the objective of equivalence to G.722 at 56/64 kbit/sec.
- 16 kbit/s is the main bit-rate. It is required the ability of the candidate to scale in bit rate to lower bit-rates (less then 16 kbit/s) and up to 24 kbit/s with no fundamental changes in either the technology or the algorithm used.
- Robustness to frame erasures and random bit errors.
- Low algorithmic delay (frame size of 20ms or integer sub-multiples)

- Applications

In the following the applications foreseen for a wideband (7 kHz bandwidth) speech coder around 16 kbit/s are listed:

1. Voice over IP (VoIP) and Internet Applications

Features: Wideband transmission over the internet
High quality speech for IP video-conferencing
Robust under background noise conditions
Robust to frame erasures
Scalable bit rate (e.g. 12,...,24 kbit/s)
Main focus on speech, good music performance at higher bit rates desirable

2. Mobile Communications

Features: High-quality speech for third generation services
Mainly to be used under relatively small residualchannel error conditions
Robust under background noise conditions
Robust to random bit errors
Robust to frame erasures
Scalable bit rate (e.g. 12,...,24 kbit/s)
Main focus on speech

3. PSTN applications

Features: High-quality audio-conferencing

Business applications (point-to-point, multi-point-links)

Robust under background noise conditions

Fixed bit rate sufficient

Main focus on speech

4. ISDN wideband telephony

Features: High-quality audio conferencing for multi-point applications

Robust to background noise

Fixed bit rate sufficient

Main focus on speech

5. ISDN videotelephony and video-conferencing

Features: Enabling high-quality speech using only one ISDN channel

Improve video quality using two or more ISDN channels

Robust under background noise conditions

Fixed bit rate sufficient

Main focus on speech

The experts group consider that VoIP, Internet Applications and Mobile Communications (listed in item 1 and 2 of the application list) are the primary applications. The Terms of Reference reflect that consideration.

The performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s are shown in Table 1.

Table 1 – (Part 1/7)

Performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s (Note 1)

Parameter	Requirement	Objective
Bit-rate(s)	• around 12 kbit/s (Note 2) • not exceeding 16 kbit/s (Note 3) • around 20 kbit/s (Note 2) • not exceeding 24 kbit/s (Note 3)	Scalable on finer increments of bit-rate. The number of bit per frame should be an integer multiple of 8.
Speech (single speaker) in error-free condition at input signal nominal level -26 dB with respect to the overload point (Note 4):		
1) at around 12 kbit/s (Note 5)	No requirement	Not worse than ITU-T Rec. G.722 at 48 kbit/s
2) at 16 kbit/s	Better than ITU-T Rec. G.722 at 48 kbit/s	Not worse than ITU-T Rec. G.722 at 56 kbit/s
3) at around 20 kbit/s (Note 5)	No requirement	Better than the 16 kbit/s under the same condition
4) at 24 kbit/s	Not worse than ITU-T Rec. G.722 at 56 kbit/s	Not worse than ITU-T Rec. G.722 at 64 kbit/s
Narrowband speech (single speaker) in error-free condition at input signal nominal level -26 dB with respect to the overload point: (Note 6)		
1) at around 12 kbit/s	No requirement	
2) at 16 kbit/s	Not worse than ITU-T Rec. G.726 at 32 kbit/s (Note 7)	
3) at around 20 kbit/s	No requirement	
4) at 24 kbit/s	Not worse than ITU-T Rec. G.726 at 32 kbit/s (Note 7)	

Table 1 – (Part 2/7)

Performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s

Speech (single speaker) in error condition at input signal nominal level –26 dB with respect to the overload point:		
• Robustness to random bit errors (BER=10^{-3})	t.b.d. (Note 8)	t.b.d.
• Detected frame erasures (3% Random) (Note 9, 10)		
1) at around 12 kbit/s	No requirement	No more than 10% additional degradation, in terms of PoW (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 48 kbit/s under error free condition
2) at 16 kbit/s	No more than 10% additional degradation, in terms of PoW (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 48 kbit/s under error free condition	No more than 10% additional degradation, in terms of PoW (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 56 kbit/s under error free condition
3) at around 20 kbit/s	No requirement	Better than the 16 kbit/s under the same condition
4) at 24 kbit/s	No more than 10% additional degradation, in terms of PoW (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 56 kbit/s under error free condition	No more than 10% additional degradation, in terms of PoW (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 64 kbit/s under error free condition

Table 1 – (Part 3/7)

Performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s

Speech (single speaker) quality dependency on the input signal level between –36 dB and –16 dB with respect to the overload point :		
Low level input speech (-36 dBov)		
1) at around 12 kbit/s	No requirement	Not worse than ITU-T Rec. G.722 at 48 kbit/s with –36 dBov input level
2) at 16 kbit/s	Better than ITU-T Rec. G.722 at 48 kbit/s with –36 dBov input level	Not worse than ITU-T Rec. G.722 at 56 kbit/s with –36 dBov input level
3) at around 20 kbit/s	No requirement	Better than the 16 kbit/s under the same condition
4) at 24 kbit/s	Not worse than ITU-T Rec. G.722 at 56 kbit/s with –36 dBov input level	Not worse than ITU-T Rec. G.722 at 64 kbit/s with –36 dBov input level
High level input speech (-16 dBov)		
1) at around 12 kbit/s	No requirement	Not worse than ITU-T Rec. G.722 at 48 kbit/s with –26 dBov nominal input level
2) at 16 kbit/s	Better than ITU-T Rec. G.722 at 48 kbit/s with –26 dBov nominal input level	Not worse than ITU-T Rec. G.722 at 56 kbit/s with –26 dBov nominal input level
3) at around 20 kbit/s	No requirement	Better than the 16 kbit/s under the same condition
4) at 24 kbit/s	Not worse than ITU-T Rec. G.722 at 56 kbit/s with –26 dBov nominal input level	Not worse than ITU-T Rec. G.722 at 64 kbit/s with –26 dBov nominal input level
Speech (single speaker) quality dependency on speakers at input signal nominal level (Note 11):		
1) at around 12 kbit/s	No requirement	Not worse than ITU-T Rec. G.722 at 48 kbit/s
2) at 16 kbit/s	Better than ITU-T Rec. G.722 at 48 kbit/s	Not worse than ITU-T Rec. G.722 at 56 kbit/s
3) at around 20 kbit/s	No requirement	Better than the 16 kbit/s under the same condition
4) at 24 kbit/s	Not worse than ITU-T Rec. G.722 at 56 kbit/s.	Not worse than ITU-T Rec. G.722 at 64 kbit/s

Table 1 – (Part 4/7)

Performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s

Music in error-free condition at input signal nominal level –26 dB with respect to the overload point:		
1) at around 12 kbit/s	No requirement.	No requirement.
2) at 16 kbit/s	No requirement.	No requirement.
3) at around 20 kbit/s	No requirement	No requirement
4) at 24 kbit/s	No requirement	Not worse than ITU-T Rec. G.722 at 48 kbit/s
Performance in the presence of the following background noises for 1, 2, asynchronous tandem encodings (foreground signal is single speaker): • office noise (SNR=15dB) • babble noise (SNR=20dB) • car noise (SNR=15dB) • interfering talker (SNR=15dB) • *performance in the presence of reverberant speech conditions. For further study (Note 12)*		
1) at around 12 kbit/s	No requirement	No more than 10% additional annoying degradation, in terms of annoying or very annoying (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 48 kbit/s
2) at 16 kbit/s	No more than 10% additional annoying degradation, in terms of annoying or very annoying (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 48 kbit/s	No more than 10% additional annoying degradation, in terms of annoying or very annoying (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 56 kbit/s
3) at around 20 kbit/s	No requirement	Better than the 16 kbit/s under the same condition
4) at 24 kbit/s	No more than 10% additional annoying degradation, in terms of annoying or very annoying (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 56 kbit/s	No more than 10% additional annoying degradation, in terms of annoying or very annoying (i.e. % of 1+2 votes), with respect to ITU-T Rec. G.722 at 64 kbit/s

Table 1 – (Part 5/7)

Performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s

Tandeming capability for asynchronous tandemings (Note 13):		
1) at around 12 kbit/s	No requirement.	No requirement.
2) at 16 kbit/s	Not worse than ITU-T Rec. G.722 at 48 kbit/s under the same tandem condition (2 tandemings)	Not worse than ITU-T Rec. G.722 at 56 kbit/s under the same tandem condition (2 tandemings)
3) at around 20 kbit/s	No requirement.	No requirement.
4) at 24 kbit/s	Not worse than ITU-T Rec. G.722 at 56 kbit/s under the same tandem condition (2 tandemings)	Not worse than ITU-T Rec. G.722 at 64 kbit/s under the same tandem condition (2 tandemings)
Transcoding with ITU-T G.722 or other standards	For further study	For further study
One-way coder/decoder delay (Note 14, 15):		
• frame size	20 ms (or integer sub-multiples of 20ms)	10 ms (or integer sub-multiples of 10 ms)
• total codec delay (algorithmic delay+processing delay)	≤ 50 ms	≤ 25 ms
Encoder-decoder synchronization (Note 16)	Provided externally	
Capability to transmit voice-band data (Note 17)	Not needed	For further study
Capability to transmit information and signaling tones (Note 18)	DTMF	

Table 1 – (Part 6/7)

Performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s

Effect of switching signal sources to the codec (Note 19) Effect of switching between bit-rates (Note 20) Convergence time (Note 21) 1) at around 12 kbit/s 2) at 16 kbit/s 3) at around 20 kbit/s 4) at 24 kbit/s	Capability required Capability required For further study For further study For further study For further study	 No annoying effects For further study For further study For further study For further study
Idle channel noise (Note 22) • unweighted • weighted • single frequency	 Less than –66 dBm0 For further study For further study	
Sampling rate (Note 23) Timing requirements (Note 24) Jitter tolerance (Note 25)	16 kHz Equivalent to those contained in ITU-T Rec. G.722 as preliminary indication The network interface to the codec is assumed to be compliant to the jitter limits appropriate to digital equipments as specified in ITU-T Rec. G.823, G.824 and I.430 (Note 26)	
A/D and D/A converter accuracy	Testing performed with 14 bit linear PCM. 15 and 16 bit linear PCM are for further study.	
Overload point of the A/D and D/A converters	As for ITU-T Rec. G.722 as preliminary indication (+9dBm0 ±0.3dBm0)	

Table 1 – (Part 7/7)

Performance requirements and objectives for a wideband (7 kHz) speech coding algorithm around 16 kbit/s

Attenuation/frequency response of encoder and decoder analog circuitry	As for ITU-T Rec. G.722 as preliminary indication	
Nominal frequency range (Note 27)	50 to 7000 Hz	
Digital transport compatibility	ITU-T Rec. H.221, H.320, H.323, H.324	
Variable bit-rate compatibility	Capability of switching bit-rate at frame boundaries required. The frame size should be equal for the different bit-rates.	For further study
Interoperability		Interoperability with 3G and 2G systems (Note 28)
Complexity	As low as possible (Note 29). Encoder and decoder in 1 DSP (single CPU)	
Memory	As low as possible in terms of RAM used	For further study
Specification description and implementation	Bit-exact 16 bit fixed-point modular ANSI-C code (Note 30) electronic format, using ITU-T basic operators (Note 31)	Interoperable floating-point specification to follow after fixed-point specification

Notes to Table 1

1. The requirements and objectives refer to the distortion introduced between the defined input PCM interface of the coder and the defined output PCM interface of the decoder, unless otherwise specified.

2. The algorithm is requested to operate at around that bit-rate. A specific bit-rate will be defined for the **Selection phase**, taking into account the need to cover specific applications, within the scope of the listed applications, and to increase the interoperability among different systems.

3. Hereafter in the Terms of Reference, the wording “at 16 kbit/s” should be intended as “not exceeding 16 kbit/s” and the wording “at 24 kbit/s” should be intended as “not exceeding 24 kbit/s”, as applicable to the two main bit-rates of the candidate algorithm.

4. When all requirements are met, the objectives will be used as one of the criteria to discriminate between candidates.

5. That bit-rate is not tested in the Qualification test. In the **subsequent testing phases** the objectives will be considered and tested.

6. In some possible scenarios the wideband speech coder may have to code a signal coming from a narrowband source. One such application is the wideband teleconferencing scenario where some of the callers participate in the call using narrowband terminal equipment while others use the wideband option. It should therefore be possible for the wideband speech coder to encode a narrowband signal with high quality.The input characteristic of the narrowband speech is for further study.

7. The requirement for narrowband input speech is not tested in the Qualification test, where the candidate proponents are requested to provide only <u>*demonstration material*</u>. This requirement should be tested in the **Selection test**.

8. The candidate proponents are requested to provide <u>*demonstration material*</u> for the Qualification Test including the 16 and 24 kbit/s operation modes subjected to a random BER=10^{-3}. Information is also welcome on different levels of BER and on the exposure of different portion of the bit-stream to different BERs.

9. Detection of frame erasures will be indicated to the decoder by external signals. If the test codec has a frame size other than 20 ms, a 20 ms frame boundary will be used for FER testing. Hence, multiples of 20 ms will be erased in the event of frame erasures.

10. The evaluation of speech quality in detected random frame erasure conditions will consider a requirement in terms of percentages of PoW (Poor or Worse) allocated in addition to the PoW percentage obtained from the reference codec. The additional degradation is an absolute increase of 10% (or 5%), that means the %(P+B) is first determined for the reference codec and an absolute value of 10 (or 5) is added to produce the increased reference degradation.
The new 'absolute' criterion, introduced to avoid using the ΔMOS concept (subject to relative shifts depending on the context of the experiment, the instructions to the team, the listeners, etc.), needs that the limits at 95% confidence level of PoW are calculated, following a defined computation rule.

11. Further studies are required to establish an appropriate methodology to test speaker dependency.

12. Further studies are required to identify realistic reverberant conditions to be tested (room geometry, distance between microphone and talker(s), reverberation time vs. frequency, etc.). The requirements and objectives related to this type of signal are to be defined.

13. A primary application requiring tandemed codecs is in multipoint control units (MCU), where signals are also mixed from several codecs before recoding and transmission. Test conditions with 3 tandemings will be considered in the Characterization phase.

14. Algorithmic delay includes the frame size, that is the block size of acquired input signal, plus any other delays inherent in the algorithm (look ahead). Processing delay is the time taken by encoder to process samples and by decoder to decode the incoming bit stream. Total codec delay is the sum of algorithmic delay and processing delay.
It is assumed that the processing delay is equal to the frame size.
The channel transmission delay assumes that transmission occurs on a serial channel matched to the bit rate of the codec, typically this is equal to the frame size. The total system delay is the sum of algorithmic delay, processing delay, channel transmission delay and such other delays caused by the test equipment and interfaces connected to these. The total system delay can be measured. The test system delay consists of the delay caused by the test equipment and the interface between the wideband speech encoder/decoder and the test equipment. The test system delay can be measured by passing PCM data directly through the system, bypassing only wideband speech encoder and decoder. The total codec delay can be calculated by subtracting the test system and channel transmission delay from the total system delay.

15. The size of frame acquired by the encoder should be 20ms or a sub-multiple of 20ms, as a requirement. For example, xz = 20ms, where x is an integer and z is the frame size in ms. Maintaining 10ms boundaries provides easy compatibility with ITU-T Rec. H.221.

16. Bit stream alignment between encoder and decoder.

17. ITU-T Rec. H.221 provides for data transmission in ISDN applications. In other applications, including DCME, PCME and ATM networks, separate transmission facilities will be provided for voice-band data.

18. The actual distortion requirements and objectives for the tones to be transmitted are for further study.

19. In a voice activated MCU where the broadcaster is selected according to voice signal levels, there should be no annoying artifacts caused by switching signal source to the codec.

20. Switching between operating bit-rates shall not introduce subjectively noticeable effects.

21. The convergence time should be as small as possible in order not to miss any significant parts of words. The convergence time of a speech codec is the amount of time it takes for the state in the decoder to approximate the state in the encoder, when the decoder has been reset to its initial state during the middle of a talkspurt. This definition readily suggests a methodology for measuring convergence time.
Begin with a test signal of speech. It is encoded and output bit stream is saved. The entire bit stream is decoded and the output is uninterrupted decoded speech. In this encoding and decoding, the states of the encoder and decoder are initialized to be equal.
Next, remove the beginning of the bit stream. Decode the remainder of the bit stream, having re-initialized the decoder first. Receive a decoded output signal which will evolve toward the true output signal decoded earlier. Then measure the sequence of individual segment signal-to-noise ratios (SNR). The convergence time will be the time beyond which the segment SNR always stays above the pre-determined threshold.

22. Idle channel noise measured with an encoder and decoder connected back-to-back as described in ITU-T Rec. G.722. Three cases are considered:
    - unweighted (unweighted noise power measured in the frequency range 50 to 7000 Hz)
    - weighted (weighted noise power measured in the frequency range 50 to 7000 Hz)
    - single frequency (the level of any single frequency, in particular 8000 Hz, the sampling frequency and its multiples, measured selectively).

23. Nominal sampling rate of the analog-to-digital and digital-to-analog converters.

24. Accuracy of the analog-to-digital and digital-to-analog converters clocks.

25. Jitter tolerance of the incoming bit-stream (peak-to-peak variation) that avoids introduced errors at the decoder.

26. In order to ensure that any equipment can be connected to any recommended hierarchical interface within a network, it is necessary to arrange that the input ports of all equipment are capable of accommodating levels of jitter up to the maximum network limit defined in ITU-T Rec. G.823 “The control of jitter and wander within digital networks which are based on 2048 kbit/s hierarchy” and ITU-T Rec. G.824 “The control of jitter and wander within digital networks which are based on the 1544 kbit/s hierarchy”. For equipments connected to an ISDN Basic Rate Access, the NT jitter characteristics contained in ITU-T Rec. I.430 are also relevant.

27. The nominal 3 dB bandwidth measured with an encoder and decoder connected back-to-back as described in ITU-T Rec. G.722. The measurement should be performed by injection of white noise to the codec.

28. If there is one or more than one bit-rate in common with the source bit-rates for wideband coding defined in 3G (including at least 3GPP and 3GPP2) and 2G systems, the objective is to provide bit-stream compatibility at those bit-rates, i.e. the bit-stream produced by this wideband encoder should be correctly decoded by the 3G and 2G wideband decoder and viceversa.

29. When all requirements are met the complexity and memory figures will be used as one of the criteria to discriminate between candidates.

30. Modular means a software implementation made in accordance to the guidelines given in ITU-T Software Tools Library User Manual.

31. The specific version of the ITU-T basic operators to be used will be defined before starting the Selection phase.

- **Schedule for wideband coding around 16 kbit/s**

A timetable for the development of a wideband speech coding algorithm around 16 kbit/s is reported in the following:

Date	Activity
25 February 2000	Final draft of the Wideband Qualification Test Plan available on SG16 Informal FTP Area (IFA) with two annexes: Specification of candidate algorithm interface Suggested cross-checking procedure
29 February 2000	Comments received by the Q.20/16 Rapporteur (via e-mail) on the MoUs (A, B, C) and NDA. (Note: MoU(A) will be signed between each candidate proponent and the ITU-T (i.e. one per candidate). MoU(B) will be signed between the host lab and the ITU-T. MoU(C) will be signed between the listening lab and the ITU-T. The NDA requires the host and cross-checking laboratories to keep the candidate executables in confidence.)
17 March 2000	Final version of LOI, MoU (A), MoU (B), MoU (C) and NDA available on SG16 IFA. (Note: The LOI will contain the intention to submit a candidate algorithm and the intention to follow item 2.2 of the ITU-T patent policy.)
24 March 2000	LOI received by ITU-T TSB from Candidate Proponents.
28 March 2000	Call for bids (host/listening labs) issued by the Q.20/16 Rapporteur via e-mail.
18 May 2000 (end of SG12 meeting)	Final version of the Wideband Qualification Test Plan approved. Review of bids. Pricing and work allocation defined.
30 June 2000	Two signed copies of MoU(A) received by ITU-T TSB from Candidate Proponents.
7 July 2000	Two signed copies of NDA received by Host Lab and Cross-checking Lab from each Candidate Proponent (if required).
7 July 2000	Qualification costs placed in ITU-T bank account by close of business on 7 July (proof of transfer may be requested). Completed copies of NDA received by Candidate Proponents from Host Lab and Cross-checking Lab (if required).
24 July 2000	Preliminary submission of executable code to the Host Laboratory.
31 July 2000	The Host Laboratory receives the source material from the Listening Laboratory. The Host Laboratory generates the Noise Files and the Error Files.
31 July 2000	Two signed copies of MoU(B) and MoU (C) received by ITU-T TSB from Host Laboratory and Listening Laboratory, respectively.
4 August 2000	The Host Laboratory receives the (floating-point) executable code, with test vectors, from each Candidate Proponent.
7 August 2000	The Cross-checking Lab receives a subset of source material for cross-checking activities.
11 August 2000	The Cross-checking Lab receives the necessary executables for cross-checking activities.
4 September 2000	The Listening Laboratory will receive the processed (already cross-checked) material for the Reference and Calibration Systems. The Listening Laboratory will receive, with blind mapping, the processed (already cross-checked) material for the Candidate Algorithms (starting point).
22 September 2000	The Listening Laboratory will receive, with blind mapping, the processed (already cross-checked) material for the Candidate Algorithms (ending point).

18 October 2000	Deliverables (raw data and test results) are sent via e-mail by the Listening Laboratory to the Q.20/16 Rapporteur (Mr Rosario Drogo De Iacovo, CSELT/ITALY; e-mail: rosario.drogodeiacovo@cselt.it), to the WP3/16 Chairman (Mr Simão F. Campos Neto, COMSAT/USA; e-mail: simao@ctd.comsat.com) and to the appointed contact of the Candidate Proponent.
23 October 2000	A notification (e-mail) is sent by the Q.20/16 Rapporteur informing the group, through the ITU-T e-mail reflector (e-mail: tsg16q20@itu.int) that the deliverables (raw data and test results) have been received.
27 October 2000	If the Candidate Proponent wishes to withdraw their ITU-T Wideband Candidate Algorithm they must inform the Q.20/16 Rapporteur via e-mail and fax not later than close of business on 27 October.
1 November 2000	Submission of deliverables to ITU-T TSB as Delayed Contributions: • Ecah Candidate Algorithm that was not withdrawn submits a test report (in the form of a technical contribution) with its test results, as received by the Listening Laboratory. The test results will be presented according to the format described in the ITU-T Wideband Qualification Test Plan. • Each Candidate Proponent submits its deliverables (high-level description of the algorithm, complexity evaluation, IPR statement). The high-level description of the algorithm shall also include information on how the bit reduction is performed to operate at around 12 and 20 kbit/s.
3 November 2000	Demonstration material from each Candidate Proposal made available on IFA. The demonstration material, using blind mapping, is provided by the Host Laboratory on behalf of each Candidate Proponent.
Study Group meeting Nov. 13-17, 2000	At the Study Group meeting in November 13-17, 2000 : • A clear statement is made indicating whether or not the Host Laboratory and Listening Laboratory have been judged compliant in meeting their deliverables according to the contents of the Memoranda of Understanding. • Evaluation of the test results of the Qualification Phase. • Draft version of the Selection Test Plan.
	Final version of the Selection Test Plan available on SG16 IFA.
	Start the organization of the Selection Phase (based on MoUs and NDA as done in the Qualification Phase).
	Start of Host Laboratory sessions.
	End of Host Laboratory sessions.
	Start of Listening Laboratory sessions.
	End of Listening Laboratory sessions.
	Each Listening Laboratory submits a report (in the form of a technical contribution) with test results. Global analysis performed.
	Detailed description of Candidate Algorithms submitted as White Contributions.
Study Group meeting April/May, 2001	Evaluation of the results of the Selection Phase. Evaluation of the algorithm complexity. **Determination of Recommendation and C-source code.** Draft version of the Characterization Test Plan.
	Final version of the Characterization Test Plan available on SG16 IFA.
	Start the organization of the Characterization Phase (based on MoUs and NDA as done in the Qualification Phase).
	Start of Host Laboratory sessions.
	End of Host Laboratory sessions.
	Start of Listening Laboratory sessions.
	End of Listening Laboratory sessions.
	Each Listening Laboratory submits a report (in the form of a technical contribution) with test results. Global analysis performed.
Study Group meeting October/November, 2001	Evaluation of the results of the Characterization Phase. Completion of the algorithm implementation verification test procedures. **Decision of Recommendation.**