ETSI TR 102 430 V1.1.1 (2005-06)

Technical Report

Speech Processing, Transmission and Quality Aspects (STQ); Basic Issues concerning the Quality of Speech over Packet Technology (both Internet and Next Generation Networks)

Reference

DTR/STQ-00037

Keywords

noise, QoS, speech

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

Individual copies of the present document can be downloaded from:
http://www.etsi.org

The present document may be made available in more than one electronic version or in print. In any case of existing or perceived difference in contents between such versions, the reference version is the Portable Document Format (PDF). In case of dispute, the reference shall be the printing on ETSI printers of the PDF version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
http://portal.etsi.org/tb/status/status.asp

If you find errors in the present document, please send your comment to one of the following services:
http://portal.etsi.org/chaircor/ETSI_support.asp

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://webapp.etsi.org/IPR/home.asp).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Speech Processing, Transmission and Quality Aspects (STQ).

1 Scope

The present document gives an overview and critique of the current developments in speech quality in packet based networks, including both the Internet and planned Next Generation Networks. The aim is to provide an overview of the main issues and their significance and to identify areas that need further work.

The present document is written at a level suitable for technical managers and engineers who are relatively new to the subject. It does not cover the issues at the depth that will be needed by a detailed specialist but its breadth should provide a valuable perspective for the specialist.

The present document includes a discussion of how the market for telecommunications is developing and how the issues concerning speech quality in packet based networks will play a pivotal role in the increasing competition between the proposed telco NGN networks and the public Internet.

The present document attempts to answer the following questions:

- What are the main issues?
- Which issues are being solved satisfactorily?
- How useful are the various approaches?
- Can end-end quality be guaranteed?
- How vulnerable is voice quality in a packet environment?
- How do you design for good quality?
- Do we need end-end signalling or lower layer class based treatment or both?

The present document is applicable to all forms of real-time 2-way conversational speech communications over packet based networks, including both telco NGN networks and the public Internet.

2 References

For the purposes of this Technical Report (TR), the following references apply:

[1] ETSI TS 101 329-2 (V2.1.3): "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3; End-to-end Quality of Service in TIPHON systems; Part 2: Definition of speech Quality of Service (QoS) classes".

[2] ITU-T Recommendation E.164: "The international public telecommunication numbering plan".

[3] ITU-T Recommendation E.800: "Terms and definitions related to quality of service and network performance including dependability".

[4] ITU-T Recommendation G.168: "Digital network echo cancellers".

[5] ITU-T Recommendation G.711: "Pulse code modulation (PCM) of voice frequencies".

[6] ITU-T Recommendation G.723.1: "Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s".

[7] ITU-T Recommendation G.726: "40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM)".

[8] ITU-T Recommendation G.729A: "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)".

[9] ITU-T Recommendation G.1000: "Communications Quality of Service: A framework and definitions".

[10] ITU-T Recommendation H.323: "Packet-based multimedia communications systems".

[11] ITU-T Recommendation P.64: "Determination of sensitivity/frequency characteristics of local telephone systems".

[12] ITU-T Recommendation P.501: "Test signals for use in telephonometry".

[13] ITU-T Recommendation P.862: "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs".

[14] ITU-T Recommendation Y.1540: "Internet protocol data communication service - IP packet transfer and availability performance parameters".

[15] ITU-T Recommendation Y.1541: "Network performance objectives for IP-based services".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

3,1 kHz handset telephony: real-time two-way speech communication within the frequency range approximately from 300 Hz to 3 400 Hz using one or more telecommunication networks with suitable terminal equipment connected to the network termination points, characterized by:

- presentation of an acoustical speech signal to the mouthpiece of a traditionally shaped handset:
  - either analogue transport of said speech signal under real-time conditions through and by telecommunication networks: said networks being intended for telephony applications between network termination points; or
  - filtering of said speech signal to the frequency range approximately from 300 Hz to 3 400 Hz; transformation of said speech signal either by waveform or by non-waveform (speech analysis) encoder; transport and processing of said speech signal under real-time conditions through and by telecommunication networks: said networks being intended for telephony applications between network termination points; back transformation (speech synthesis) of said speech signal by the respective decoder;
- acoustical presentation of said speech signal in the frequency range approximately from 300 Hz to 3 400 Hz by the earpiece of a traditionally shaped handset.

instant messaging service: proprietary service that provides peer-peer communications over the Internet and may also provide the ability to make calls to the PSTN or to receive calls from the PSTN

NOTE: Most instant messaging services provide presence indications for the users on the subscriber's contact list.

quality of service: collective effect of service performances which determine the degree of satisfaction of a user of a service. It is characterized by the combined aspects of performance factors applicable to all services, such as:

- service operability performance;
- service accessibility performance;
- service retainability performance;
- service integrity performance;
- other factors specific to each service.

NOTE: From ITU-T Recommendations E.800 [3] and G.1000 [9].

speech quality: expression of the degree of customer satisfaction with conversational speech transmission

NOTE: Speech quality depends on the quality of the whole speech path from the talker at one end of the connection to the listener at the other, and can be categorized into two types of quality: quality which is mainly dependent on handset acoustics and quality which is mainly dependent on the transmission medium. Telecommunications services where special attention needs to be paid to speech quality, such as audio teleconferencing and voice mail, should also be considered.

telco: telecommunications organization (e.g. company, operator, provider)

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

3GPP	3rd Generation Partnership Project
ADSL	Asymmetric Digital Subscriber Line/Loop
API	Applications Programming Interface
ATM	Asynchronous Transfer Mode
AVT	Audio/Visual Transport
BICC	Bearer Independent Call Control
DiffServ	Differentiated Service
GSM	Global System for Mobile communication
IETF	Internet Engineering Task Force
iLBC	internet Low Bitrate Codec
IN	Intelligent Network
IP	Internet Protocol
IPDV	IP Packet Delay variation
IPER	IP packet error ratio
IPLR	IP Packet Loss Ratio
IPTD	IP Packet Transfer Delay
ISDN	Integrated Services Digital Network
ISP	Internet Service Provider
ISUP	Integrated Services User Part
ITU	International Telecommunication Union
MPLS	Multi-Protocol Label Switching
NAT	Network Address Translator
NGN	Next Generation Network
NNI	Network-Node Interface
PCM	Pulse Code Modulation
PSTN	Public Switched Telephone Network
QoS	Quality of Service
QSIG	Q reference point Signalling
RSVP	Resource reSerVation Protocol
RTP	Real Time Protocol
SIP	Session Initialization Protocol
UNI	User Network Interface
VoIP	Voice over Internet Protocol
VPN	Virtual Private Network
WLAN	Wireless Local Area Network
xDSL	(Undefined) Digital Subscriber Line/Loop

4 The Market

4.1 Overview

Overall the telecommunications market currently lacks direction. After a period of diverse investments, many of which have not been profitable, the top priority for many operators is to manage their short term financial objectives.

In terms of fundamental resources:

- local physical infrastructure remains expensive;
- transmission costs has fallen and is falling very rapidly thanks to a combination of absolute costs (cost per bit) and coding that enables more use to be made of a bit, i.e. more information may be transmitted via the same amount of bits (e.g. voice coding);

NOTE 1: Transmission bit (not information bit).

- switching costs are falling but faster for packet-based switching than for circuit-based switching, because packet based technology is benefiting from large economies of scale from its use in corporate networks and the public Internet?
- billing costs are falling only slowly, and the overheads of running a telecommunications business are increasing as a result of increased regulatory compliance costs including areas such as data protection.

This situation leaves telcos re-focusing on core business and in particular developing broadband access networks.

There is a great deal of discussion and confusion about how telecommunications networks will develop at a technical level in the next few years. Three years ago everyone was expecting the rapid and near universal adoption of IP technology but since then the whole investment climate has changed and the current situation is much less clear. Some established operators are formulating plans for a fairly rapid migration to packet technology whilst others are waiting and planning to extend the lifetime of their existing switched infrastructure for as long as possible.

In the Internet world, there is substantial growth in VoIP services, the most well known being Skype.

Figure 1 shows a credible view of how networks will change. The diagram (see note) is best viewed in colour as the colours are significant. The blue rectangle covering the whole diagram illustrates the dependence on a common IP based transmission platform, the exception being the lop left hand corner where circuit switching is still used.

NOTE 2: Adapted from a diagram produced by Mr Nozsek of Deutsche Telekom.

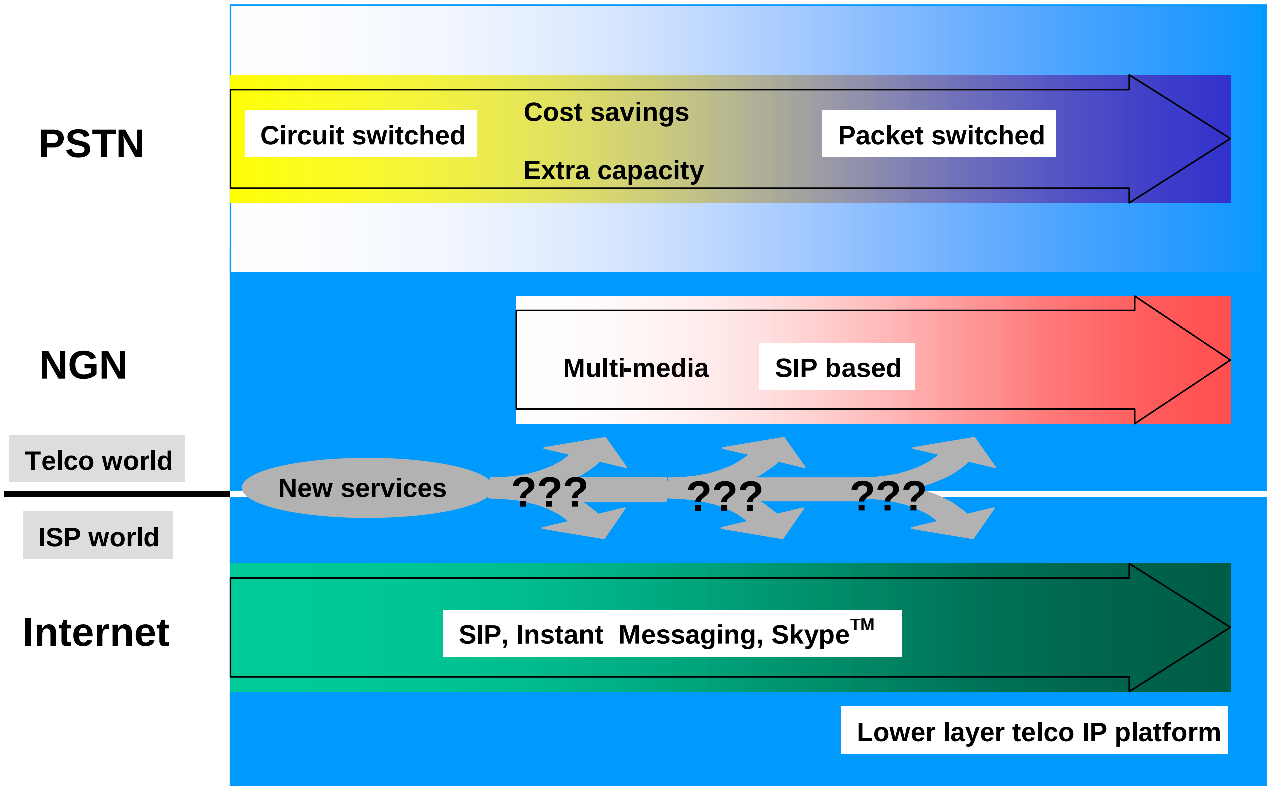


Figure 1: Network developments

4.1.1 PSTN

The PSTN/ISDN is largely circuit switched. Public services currently offered on the PSTN using E.164 numbering and G.711 PCM coding are likely to continue largely unchanged for the indefinite future because they work well and are universal.

The main justification for replacing circuit switched with packet switched networks (either IP or ATM based) is cost saving primarily in operational expenditure since the capital expenditure has already been made. This justification will grow gradually if manufacturers fail to supply adequate spares and if the expertize for software modifications is dissipated, but relatively few modifications will be needed and the current circuit switches could remain in use for at least another 10 years, at least at the local level.

There are differing reports on the current scope for justifying replacement based on savings. One operator has recently announced widespread replacement of its transit and international network based on savings in operational expenditure mostly by reducing the number of switches. Other operators have been unable to justify any change. It is also not entirely clear whether a packet based network, especially an IP based one, can reproduce all the features of the current circuit switched PSTN; a reduction in the range of supplementary services may be involved.

Where extra capacity is needed, it is less likely that operators will buy new circuit switches and some manufactures may no longer be able to supply them, so they will buy softswitches instead. Since PSTN/ISDN traffic is static or falling (except for Internet access and some calls to non-geographic numbers) there should not be too much need for additional capacity.

The solution to the growing Internet access traffic is to introduce xDSL e.g. as ADSL and so remove this growing traffic from the local switches. Some telcos are now pushing ADSL very actively and high penetration rates are being achieved in some countries.

Where circuit switches are replaced by softswitches, the aim is to make the PSTN services appear not to be changed. Thus the simplest solution is to implement the No7 Signalling Protocols over IP with minimum changes. This means using ISUP or BICC or H.323 over IP, but not SIP. Manufacturers already are doing this for transit level switches (Class 4) but few if any have yet developed soft switches with the full capability of local exchange circuit switches, but this will change.

At the international level, there are now several IP based networks that handle international traffic including traffic from incumbents and the entry into the market of these networks has helped to create an active market in international call minutes. Some of this traffic is handled on dedicated IP networks and some on the public Internet.

In summary the PSTN/ISDN service is likely to stay largely unchanged with slowly declining traffic volumes but the technology will see a slow migration to softswitches. The PSTN arrow in the diagram shows the migration by a gradual change in colour.

4.1.2 NGN

"NGN" is used here to describe the telcos" attempt to develop an IP-based platform for future services. They are particularly keen to develop a multi-service platform capable of supporting multi-media services and wish to allow separation between service or application creation and basic transport. The NGN concept continues the telco approach of "closed networks" where charges are mostly usage based.

SIP is currently the favourite protocol for these developments and work on SIP is being undertaken in 3GPP for its IP Multimedia Platform.

One of the problems with NGN is that few people have clear ideas of what services will be needed. This is one reason why the manufacturers are pursuing an "open services environment", as no one is very sure about what to do. In general the telcos say that they want to pursue technical competition in service creation, rather than standardization and they are resisting suggestions of service standardization in ETSI.

A vital question for NGN is whether it needs to include PSTN functionality, or can develop separately and in parallel to the slow migration of the PSTN to IP. This is important because it determines the extent to which the NGN needs to take account of the special features of the PSTN. If the two will develop separately and in parallel then the NGN will not need to embrace the PSTN, although both will be supported on the same underlying infrastructure. This will simplify greatly the development of the NGN. The NGN programme in ETSI sees support of the PSTN as its initial objective and the support of new multi-media services as the second objective.

Figure 2 shows the NGN arrow growing from nothing to indicate its gradual implementation.

4.1.3 Corporate VPNs

This is currently the area where telco IP-based services are growing most rapidly. The VPNs provide:

- internal voice communications;
- external PSTN access;
- services exclusive to the customer that relate to their operations;
- Internet access.

The needs for corporate and public telephony are similar technically to the provision of public telephony over IP, however the protocol is likely to be QSIG over IP since it will be necessary to provide a smooth transition for services from circuit switching to IP.

There is as yet little information about new NGN services other than ones that relate specifically to corporate operations.

Continued expansion of the VPN market is likely and it is also likely that there will be a demand for interconnection between the VPNs of different organizations. However this interconnection will only be of value where the "services" of both networks are similar at a technical level. This should be achievable for standardized services such as public telephony and its private counterpart. Where new NGN services have been developed such as video-telephony, interconnection between VPNs run by different telcos will be possible only if the "service" is similar at a technical level, which implies standardization.

4.1.4 Internet

The public Internet is the third area of development. It is by definition an open services environment but the commercial arrangements are quite different from those of the NGN because the Internet provides a global platform with access paid largely by subscription.

The range of services available on the Internet is increasing and users are able to obtain services, including voice communications, at low or zero marginal price on the Internet that previously they had to pay usage based charges for to the telcos. Thus the fixed telcos are facing a steady migration of traffic away towards the Internet and also to mobile networks.

The Internet model represents a major threat to the traffic related revenue of the telcos because the marginal costs of using the Internet are low or zero for many users who have already obtained the necessary equipment such as a PC.

Figure 2 compares the "closed" telco networks with the "open" Internet.

Current telco networks	Telco NGN networks - closed	Internet - open
• Circuit switched technology • User-user services centrally controlled by provider of transport service • Usage related charges and quality control • Access control for users and interconnection • Intelligent network/Dumb terminals • Interconnection is service related and controlled • Few/no third party services	• ATM/IP based technology • User-user services centrally controlled, with much greater scope for third party services run via APIs • Usage related charges and quality control • Access control for users and interconnection • Migration of intelligence towards the terminals • Interconnection may occur at various levels. Above the IP level it is likely to be service related and controlled	• IP based technology • No service creation - services and applications run from edge • User-user services run by users themselves • Client-host services run by independent hosts at edge • Access control for users but otherwise open • Intelligent network/Dumb terminals • Interconnection is open and only at IP level • No usage-related charges and little quality control • Gateways to telco networks have control and charging

Figure 2: Comparison of telco networks and the Internet

The distinctions are illustrated in figure 3, which compares the telco concept of the next generation network with the Internet.

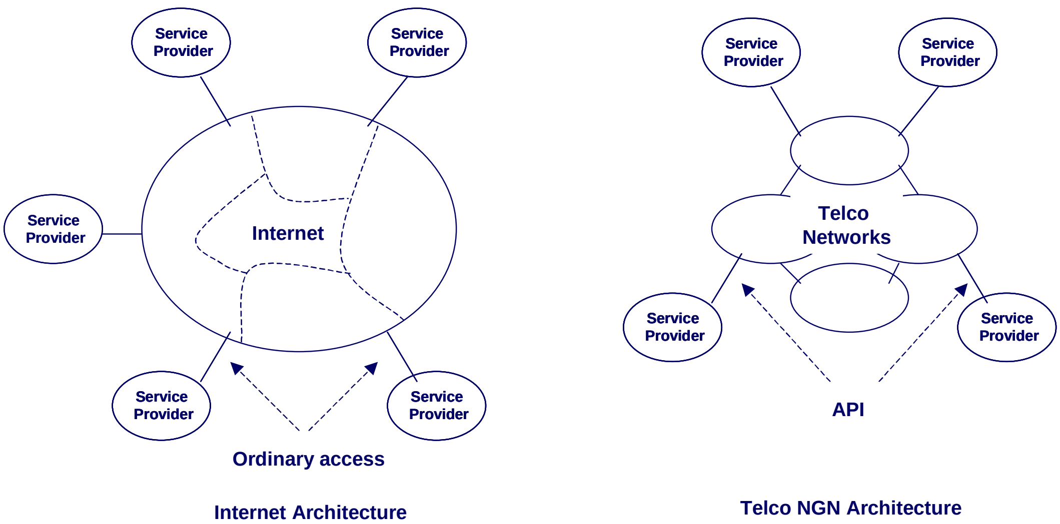


Figure 3: Comparison of telco NGN architecture and the Internet

The critical differences lie in:

- third party service provision, which can be carried out anywhere on the Internet using a normal access but require a special application programming interface on the NGN;
- interconnection, which is transparent to services in the Internet and so any service provider can offer services to anyone in the world. In contrast in the NGN the interconnection arrangements are expected to be service aware because of the need for service or "media pipe" related charging and the need for control of session border controllers. This means that a service provider will need to make prior arrangements with each network operator over whose infrastructure the service is to be offered.

4.1.5 Voice traffic - Migration towards Internet

The telcos are heavily involved in the support of the Internet in that they supply the basic transmission facilities and dial-up access and in many cases also have large businesses as ISPs, and therefore the growth of the Internet is not wholly a commercial threat. Their main risk, however, is the loss of revenue from usage based voice traffic, which is typically some three times that for access line rental.

Voice traffic can be subdivided into three categories:

- repeat calls to same people (family, friends, colleagues). This is the largest category and the one best suited to Instant Messenger services;
- calls to Government, shops, services, schools. This will be a major application for click to talk services as call centres develop Internet access;
- "random other calls", these calls are likely to remain served by the PSTN.

Figure 4 shows where voice traffic that has hitherto been carried on the fixed PSTN is migrating. The migrations are:

- slow but accelerating substitution by mobile networks. An increasing number of customers no longer bother to have fixed lines and rely wholly on mobiles;
- substitution of some short non-urgent calls by text messaging;
- substitution of various types of calls by email and Internet transactions (e.g. checking a bank statement and booking a flight);
- a slow substitutionary migration of traffic to the public Internet and corporate VPNs. This traffic is mainly frequent calls between the same small group of people (e.g. teams at work or distant family members) and calls to organizations where "click-to-talk" is provided from their web pages.

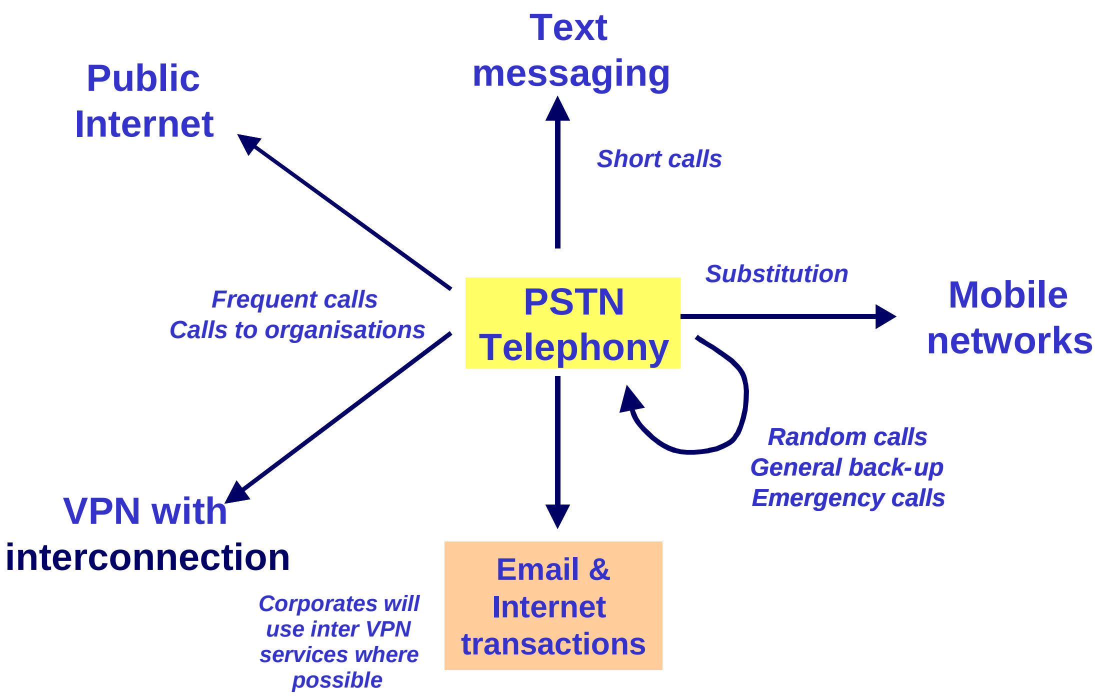


Figure 4: Migration of voice traffic

Figure 5 shows the differences in market pressure between the telcos and the Internet. The main pressure on the telcos is to reduce price of usage based services, the main pressure on the Internet world is to increase quality.

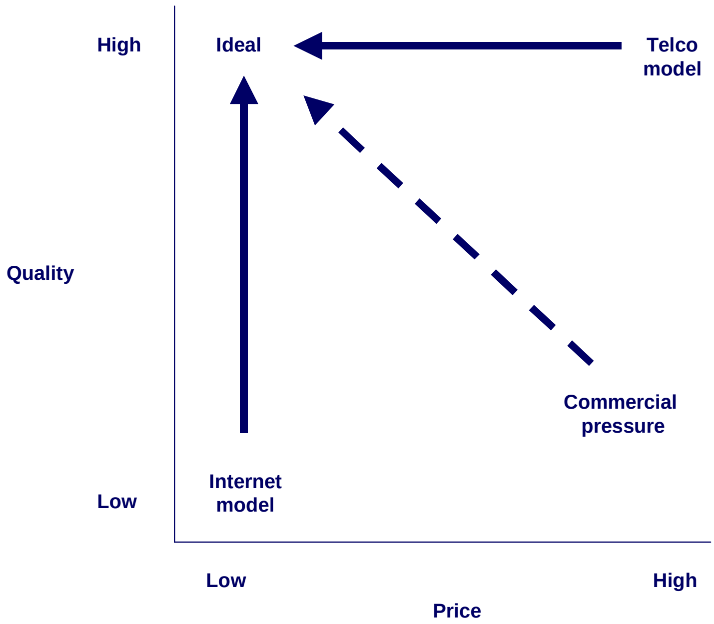


Figure 5: Market pressures

Three main issues are slowing the migration of voice traffic to the public Internet:

- transmission quality;
- ease of use;
- blocking by NATs and firewalls.

At present most of the codecs used in VoIP were designed for circuit switched applications and are badly affected by the packet loss that occurs on congested IP-based networks, however new codecs designed to tolerate packet loss are becoming available and are expected to provide adequate quality even over the Internet.

Ease of use will then remain the critical factor. Voice communications over the Internet at present depend on services such as Instant Messenger to overcome the problem of dynamic assignment of IP addresses. This creates two problems:

- there are different proprietary solutions (e.g. Microsoft, AOL and Yahoo) which results in users having to register with multiple systems, which is not popular (interestingly this is the same problem that will be created by the competition in services that the telcos now seem to favour);
- the call set-up arrangements of Instant Messaging are not always quite as simple as making an ordinary telephone call.

Dynamic assignment of IP addresses is likely to remain common for the next few years although it might reduce if there is rapid adoption of IPv6, but this seems unlikely.

In most cases, voice communications over the Internet are blocked by firewalls and Network Address Translators (NATs). Some of the causes of blocking are can be solved by changing the policy of the relevant IT Departments, but voice communications can have difficulty in traversing NATs because the NAT cannot be made to translate IP addresses and port numbers for the media streams as well as the signalling. Substantial progress has been made in this area recently for most domestic NATs and firewalls but problems remain for with the most sophisticated corporate ones.

Other developments that facilitate the migration of voice to the Internet are:

- the growing popularity of broadband Internet access with always-on capabilities. Ironically this means that if telcos accelerate the roll-out of broadband access they may facilitate the loss of some voice traffic revenue;
- growth in the use of LANs in the home, whether wireless LANs or hardwired ones. LANs are being sold in some DIY stores in some countries.

The main development that is likely to deter the migration of voice to the Internet is the introduction of flat rate tariffs by the telcos. Such tariffs are becoming more common, whether for the whole day or just for off-peak times, and they remove the cost saving incentive of using the Internet. Users seem to like flat rate tariffs because they are less vulnerable to unexpectedly high bills. Flat rate tariffs also help the telcos to reduce their costs in handling customer complaints.

The overall conclusion is that voice traffic, which has limited potential growth capability within Europe, will continue to migrate away from the fixed networks to mobile networks and to VPNs and the Internet. This migration of voice traffic is unlikely to reduce the demand for fixed access including access to the PSTN greatly as most smaller premises will require Internet access via ADSL or newer technologies and most users will wish to continue to have access to public telephony both for any-any connectivity and for use when other forms of communication fail.

4.2 Competition between NGN and the Internet

4.2.1 NGN Developments

Whilst there is a clear case for migrating private and corporate networks to an IP platform to provide integrated voice and data, there is not a clear economic case for doing so for public networks. Several operators have undertaken studies of the economic benefits of replacing circuit switched networks with IP based NGNs but have found that the benefits do not outweigh the costs, nevertheless other operators believe that they do and are embarking on replacement programmes.

The fundamental problem for fixed network operators is that traffic levels are flat or decreasing slightly for almost all traffic other than dial-up Internet traffic. The strategy of removing Internet access traffic as early as possible onto a separate network platform and leaving the circuit switched network in place therefore seems increasingly attractive and is likely to remain attractive until the maintenance costs of the circuit switches and concentrators becomes too high. This problem may occur earlier than "necessary" since many manufacturers have ceased, perhaps prematurely, manufacturing spares for this technology.

Notwithstanding this, it is unlikely that a clear universal case will emerge for replacing circuit switches within the next 4-5 years. IP based infrastructure may therefore be rolled-out in parallel as an overlay network to serve:

- new developments;
- areas where high population growth cannot be served by the existing switches;
- customers who specifically need NGNs.

Two of the hopes of the telcos are that:

- users will want to continue to have "guaranteed quality";
- service providers will pay to host services on the new telco NGN platforms.

It is not clear whether these hopes will come to fruition. Adequate quality for a high proportion of cases may prove sufficient for most customers, and innovators of new services may prefer to use the Internet and gain global reach to prospective customers at the price of basic access rather than enter special arrangements with telcos whose history of helping third party service development in the IN era was disappointing.

4.2.2 Internet Developments

The critical question for the Internet is whether quality will continue to increase or whether it is currently at its peak, due to excessive "dot-com" investments, and will deteriorate in the future. The trend for increasing dependence on the Internet suggests that people will if necessary be willing to pay more overall for Internet access, and so quality can be sustained or improved. In practice it seems that a large proportion of the costs are in the access arrangements and several countries are seeing quite high levels of demand for ADSL access, which indicates willingness to pay more for better quality. The fact that most of the bottlenecks are in the access means that it should be possible to achieve a fairly direct relationship between subscription levels and quality, giving the right economic signals to the market.

One of the main methods to improve Internet quality is to segregate traffic of different types (packet length and delay sensitivity) onto different virtual networks so that they queue separately for routers and some priority can be given to delay sensitive traffic. Techniques for such segregation have been developed (e.g. diffserv) and may be introduced in the future.

4.2.3 Hybrid developments

There is a great deal of activity in hybrid PSTN-Internet services mostly from smaller new entrant operators. The main businesses established so far are Internet based services for PSTN break-out that enable users with Internet access to make long distance and international phone calls at reduced rates, especially into countries with high termination rates. Some of these services are beginning to offer an incoming call capabilities and some regulators have allocated number blocks for their subscribers.

Other potential developments are linkages between ISPs and local exchanges so that Internet users on dial-up access can be warned on incoming telephone calls and either clear to receive then or receive them on their Internet access.

In general the hybrid developments are either specialist services or short term bypass services that will decline when better Internet access is available for more people.

4.3 NGN Services

The telcos and their suppliers who are supporting NGN developments, whether fixed or mobile, are planning to promote technical competition in the development of new services rather than the standardization of new services. Figure 6 shows the architecture planned. This approach applies to both mobile (e.g. 3GPP IP Multimedia) and fixed networks (e.g. TISPAN).

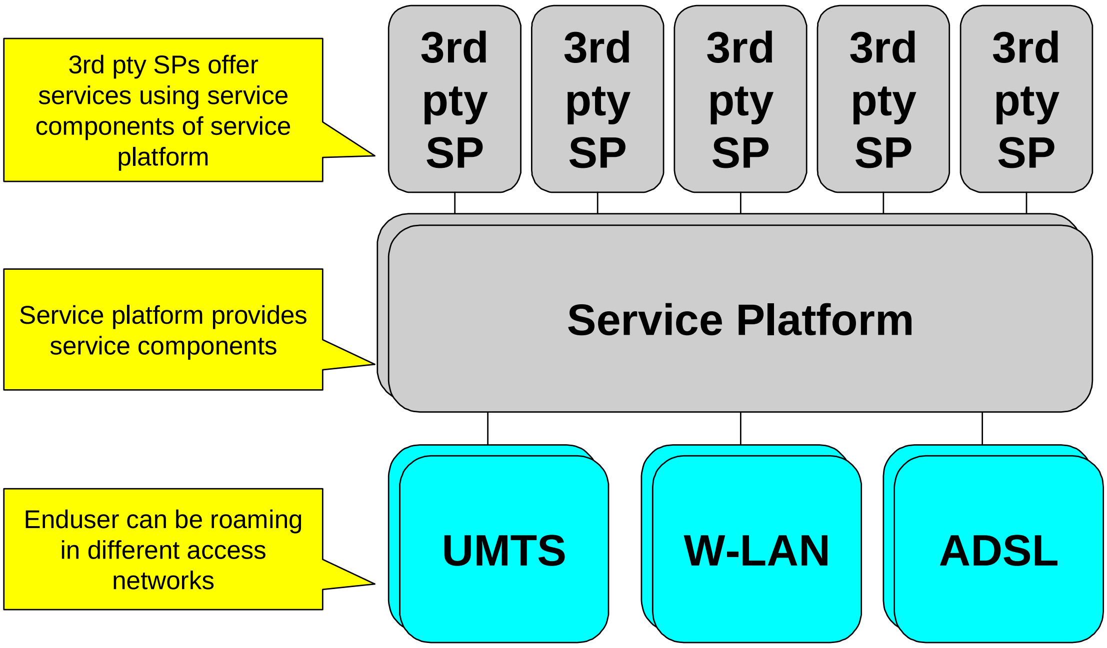


Figure 6: Architecture of NGNs

The intention is that the network operators will provide a general purpose service platform for the creation of innovative services by themselves and third parties, and that the service platforms and third party service providers will be able to charge customers on a usage basis.

The service providers will innovate in service creation and place contracts with service platforms for connectivity. This approach raises several issues:

- network operators were unwilling to promote third party service creation in the ISDN-IN era and have not yet demonstrated in practice a willingness not to favour their own vertically integrated services;
- service providers will have to negotiate connectivity agreements with the platforms of many different operators if they are to have wide coverage for their services and to have usage based billing;
- where new client-client services are provided, communications will only be possible between the customers of the same service provider unless different service providers cooperate to offer the same technical service.

It is far from clear how these developments will work out. There is a huge advantage in having a standardized service with standardized UNI and NNI interfaces for public services and also for any "private services" that could be interconnected on VPNs. The standardized UNI interface creates a large independent terminal market, and the standardized NNI provides easy any-any interconnectivity between the customers of different service providers and facilitates the development of comparable Reference Interconnection Offers. Standardization of these interfaces does not inhibit the development of new features that exist wholly within a terminal or wholly within a network.

The success of competitive service innovation compared to the standardization route will depend on:

- the extent to which better technical characteristics in a particular service influence the choice of service provider when most customers take many or all services from the same provider;
- the extent to which customers find that the loss of an any-any capability is a disadvantage when communications are possible only between customers of the same service provider. In other words, how well do the informal groups whose communications account for probably the majority of each person's communications map to the choice of service provider?
- the effect of competition from similar services on the Internet which may not have the same constraints.

Figure 7 shows some of the possible developments and their dependence on the main key issues.

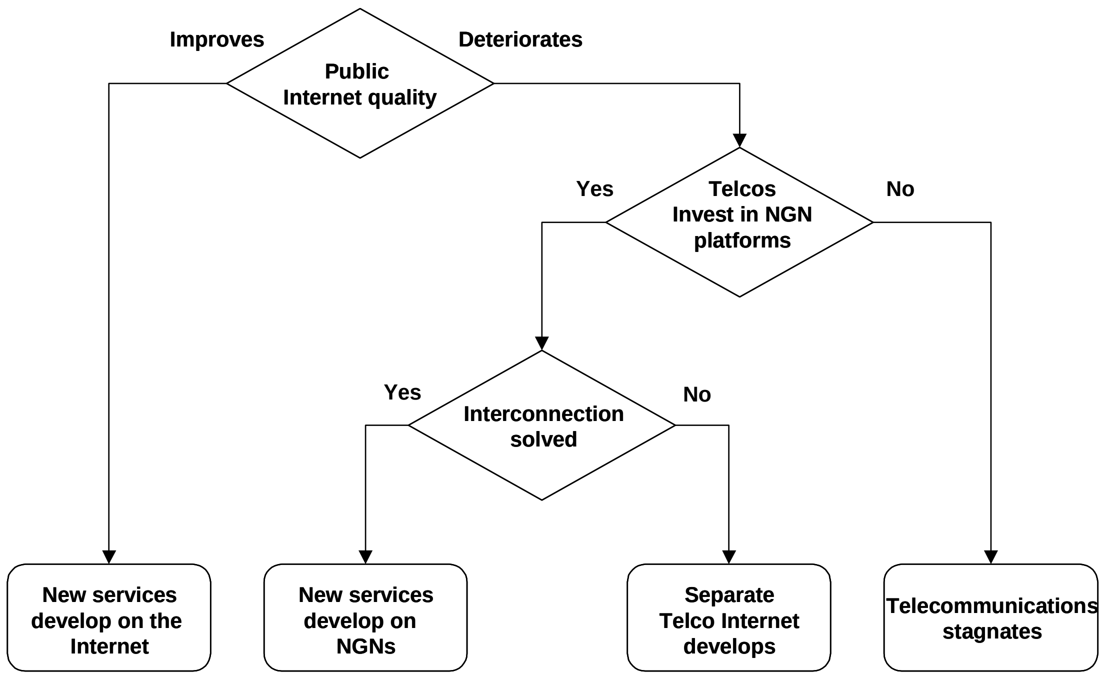


Figure 7: Possible developments

Probably the main issue is whether the quality of the Internet improves or deteriorates. If it improves then service innovation is likely to take place on the Internet. If it deteriorates, then the telcos will have more incentive to invest in NGN platforms. If the new services start to develop on the NGN platforms then the main issues will be coverage and coverage and interconnectivity. If they are solved then the current telco NGN model will prevail. If they are not solved then the telcos may have to offer an open platform of higher quality than the public Internet, i.e. an "Internet Mark 2" with higher access charges and higher charges for attaching service but without usage based charging to simplify interconnectivity.

4.4 Conclusion

Developments concerning the quality of real-time two-way communications, especially voice, over IP based networks will probably be the main factor in determining the direction of the future market. If transmission quality over the Internet proves adequate and this form of communications becomes sufficiently easy to use, then the Internet will provide very strong competition for the telco NGNs. Thus understanding the quality issues is central to understanding the future development of the market, and the development of better quality that the Internet is the key to success for the telcos.

5 Quality of Service and Network Performance

The distinction between Quality of Service and Network Performance is of central importance. Although the terms are often used imprecisely the present document uses them with the following distinct meanings:

- Quality of Service applies to the end-end service delivered to the end use and so for speech is hence a measure of mouth to ear acoustic quality. Its main parameters are:
    - distortion;
    - delay;
    - echo;
    - loudness.

- network performance applies to the performance of a network and in the case of NGNs to the performance from UNI to UNI. Its main parameters are:
  - packet loss;
  - delay;
  - delay jitter.

and their values vary with time during a call.

Figure 8 shows their relationships for a basic end-to-end speech connection with two terminals interconnected via access and one or more transit networks.

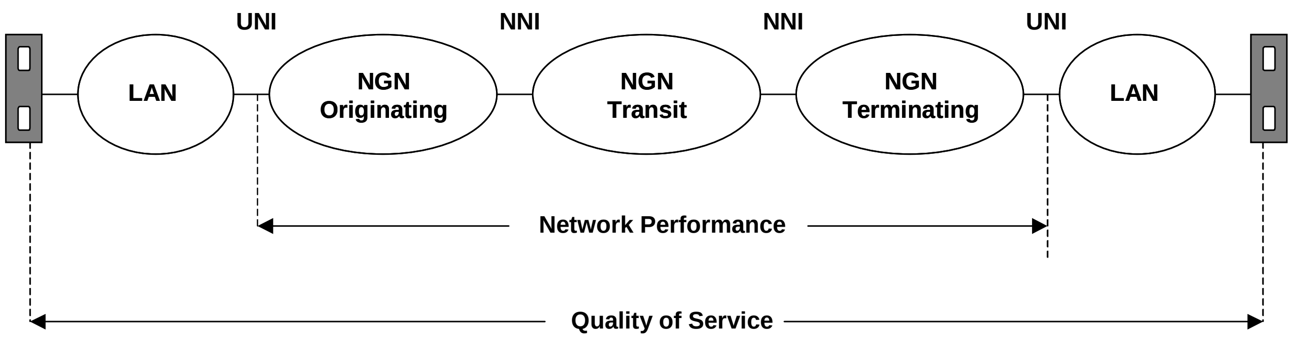


Figure 8: Possible developments

The end-to-end speech quality of the connection is affected by various transmission impairments that depend on the transmission and switching technology used and affect the end-end performance in different ways.

The overall objective of QoS work is to achieve an end-end quality that is satisfactory for the user. This is not at all a simple issue because:

- user's requirements and views on what they consider to be satisfactory differ especially in terms of their past experience and what they are paying for their service;
- no one party has responsibility for the complete end-end connection (mouth to ear) nor control over the whole connection since the terminals are normally owned by the user without control being exercised by the network operator. Furthermore whilst users may choose their own network operator, i.e. access and perhaps nearest transit operator, calls that they make will terminate on networks chosen by the called user and not by the caller itself, and neither the caller nor their service provider has any control over the terminating network and any quality choices that the called party may have made that affect these networks;
- the quality of service perceived by the user, is influenced very strongly by the terminal and in particular by the algorithms implemented in the terminal and by the way that the terminal is set-up. Furthermore if the network performance is less close to perfect, there is a very strong interaction between the terminal algorithms and the network - with the right choice of algorithms good quality can be obtained with relatively poor network performance but with the wrong choice the performance may be very poor.

The historical approach for circuit switched networks was to take a single common approach to quality based on limited bandwidth 3,1 kHz handset telephony with the aim of achieving a high degree of probability that the end to end quality would be adequate. This involved networks exercizing control over the terminals of users through supplying these terminals themselves or controlling their quality through type approval. Restrictive rules based on apportionment were applied to network design to achieve a high probability that impairments that summed such as loss in analogue networks or delay in both analogue and digital networks would in almost all practical configurations remain within an acceptable level.

Circuit switched technology was conducive to this approach because its performance was constant and independent of the network's traffic level, provided a connection could be obtained. The nature of the technology also led to a single bit rate of 64 kbit/s for the switches and this meant that deviation from this single common approach was largely impracticable. Both loss and delay in analogue networks depended on distance and the greatest distances occurred in the public networks giving the telcos control over the major sources of impairments even when terminals were liberalized. Digitalization removed the problems of loss, leaving delay as the main impairment that was subject to network planning. Delay contributes to echo, but the reducing cost of echo cancellers has helped to keep echo to tolerable levels. Delay is also an impairment in its own right and delay has generally increased with the increasing use of digital technology but in most cases delay has stayed within acceptable levels.

The introduction of packet technology has led to a wave of new interest in quality issues. The two main changes that packet technology introduces are:

- replacement of the single rate of 64 kbit/s for switches with a spectrum of rates giving scope for new codec designs and a wide range of levels of quality;
- replacement of the constant and repeatable level of quality for a connection with a variable level of quality that depends on network traffic levels.

The dependence of quality on traffic levels is the most difficult issue and the area where there is least practical knowledge and experience. The traditional approach of apportionment can be used where the network topology is known and the impairments are constant and predictable. But in packet networks neither criterion is met.

In practice there are commonly two additional issues:

- the quality problems caused by traffic congestion may be greatest in the user's own networks and so are outside the control of the public network operators;
- packet based voice terminals are much more likely to be used from PCs whose acoustic characteristics may be intrinsically poor or suffer from being incorrectly configured, but with correct set-up and cheap hardware very good quality is possible.

Thus the commonly stated NGN objective of achieving "guaranteed Quality of Service" is not practicable and is a confused statement. The objective of achieving "guaranteed Network Performance" might be practicable if it were not for the problem of control over the terminating network. An operator may set the objective of achieving "guaranteed Network Performance" on their own network or on networks that they use under contract, and this is more practicable, but there is still need for clarification about what is meant by "guarantee" and this is discussed later.

In summary, the slogan "guaranteed Quality of Service" when analysed means "reasonably consistent network performance".

6 Impairments in packet networks

6.1 Networks

At each node in a packet network, packets are held in a queue awaiting transmission. Congestion causes longer queues than normal and so increases the transmission delay for the packets. Nodes also have limited queuing capacity and queues may overflow resulting in packet loss.

The terminals at the ends of speech circuits provide jitter buffering to smooth the play-out of packets. The jitter buffer is a store of packets awaiting processing. The store has a maximum size determined by the hardware, although the size used may be varied dynamically. The packets arrive at varying times and are extracted at regular intervals by the play-out algorithm. The effect of the jitter buffer is to convert variable delay into fixed delay. The fixed delay depends on how full the jitter buffer is. If the variable delay increases so that the buffer empties (called a jitter buffer under-run) then a packet is missed. When delays reduce, the jitter buffers will fill up and if the maximum capacity is exceeded then packets will have to be discarded, and this is called buffer over-run. The play-out algorithm may be quite sophisticated and adjust the fill of the jitter buffer so that it is filled only to the extent necessary so that the probability of an under-run is low, so that the additional delay is minimized. In order to make these adjustments, packets may have to be skipped or duplicated introducing some additional distortion. Some highly intelligent algorithms may observe the values in the packets and make adjustments only when there are gaps in the speech.

Thus end-to-end transmission delay affects the quality of interactive real time communications:

- directly in terms of the average delay plus the settings of the play-out algorithms in the jitter buffers to account for delay variations (jitter);
- indirectly through packet loss where:
  - queue capacity in individual routers is exceeded and so packets are discarded (over-run);
  - variable delay increases causing the jitter buffer to empty and resulting in missing packets (under-run).

NOTE: The packet sent is not lost but a gap appears at the receiving end and is normally filled by repeating the last packet or another method of concealment.

  - Variable delay reduces causing the jitter buffer to overflow and resulting in lost packets (over-run).

The different effects are shown in the figure 9. The diagram distinguishes the network and terminal factors. The jitter buffer operation can be set to give lower delay and higher packet loss, or higher delay and lower packet loss. Codecs introduce delay that depends on the algorithm used and the processor power. There are also various different methods for handling packet loss called packet loss concealment (also known as network equalization) and these algorithms may be implemented separately from the decoder or be integrated into the decoder.

Unlike digital circuit switched networks, packet networks are not normally synchronized to a common reference. Thus the clocks at the sending and receiving ends are likely to be running at slightly different rates causing the problem called "clock skew". Clock skew leads to overflow or underflow of playout buffers at the input to the decoder so that packets are lost in one direction and gaps occur in packet arrivals in the other direction. Both affect the decoding and contribute to distortion.

Since the connection is "4-Wire" in terms of echo, echo can arise only at the far end across the ear-mouth link, where the linkage is called terminal coupling. If the terminal coupling loss is high as with a conventional handset or high quality headset, then no echo control or cancelling may be needed. If the terminal coupling loss is low, and with a loud speaking telephone it will be very low, then echo cancelling is needed and this is normally applied at the terminal itself. For example some PC operating systems offer loud speaking settings that include echo control. The echo control procedures may increase the distortion of the speech or clip the speech.

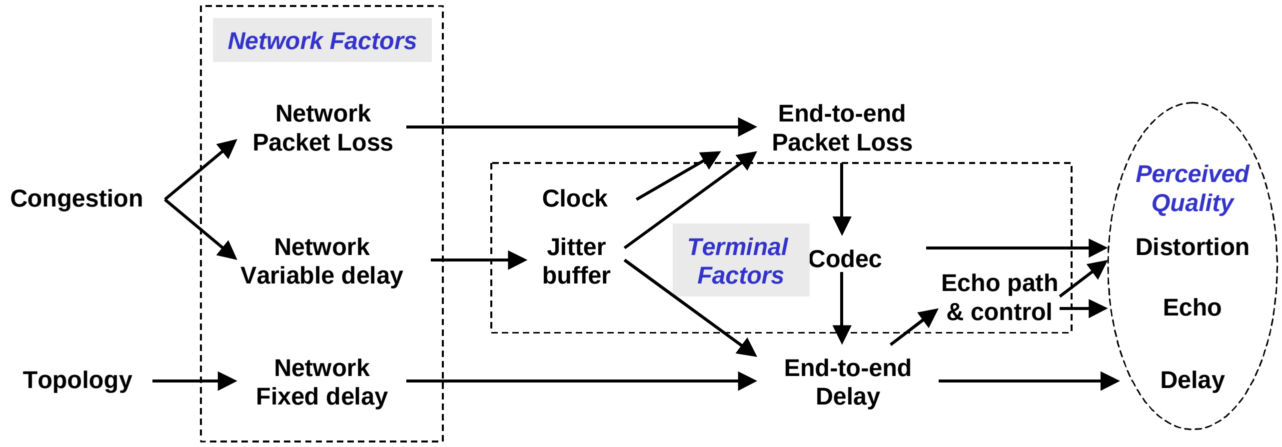


Figure 9: Impairment effects for packet based voice communications

The figure distinguishes between the effects that occur in the network and the mechanisms in the terminals that are affected and that can be used to correct for the effects in the network.

Many calls traverse multiple networks including end networks such as LANs, access networks and transit networks. End or customer networks are likely either to be wireline or wireless LANs. A wide variety of loadings is likely. Wireless LANs are especially likely to be overloaded at times and can in practice be the dominant source of quality problems for Internet access. Organizations where there is correlation between the activities of different LAN users are especially likely to experience high peaks of demand from simultaneous similar actions by users, e.g. all participants in a meeting downloading documents at the same time.

Figure 10 provides a summary of the different impairments and their sources. Those impairments that can typically degrade quality significantly are shown in bold.

	Circuit switched	Packet switched
Terminal equipment	Echo from terminal coupling Loudness rating **Delay and distortion from coding in mobile terminals**	**Echo from terminal coupling** **Distortion from packet loss** **Distortion from low bit rate coding** **Delay from jitter buffer** **Loudness rating** Delay from coding
End network e.g. LAN	None	**Congestion causing variable delay and packet loss**
Access network	None	**Congestion causing variable delay and packet loss**
Transit networks	**Delay** Distortion from transcoding	**Delay** and variable delay

Figure 10: Comparison of circuit switched and packet switched impairments

6.2 Terminal and codec issues

The analysis of impairments has shown the importance of terminal design and codecs, and their scope for compensating for impairments that arise in the networks. Terminals are especially important where the public Internet is being used and the general level of impairments may be higher than in generously dimensioned networks run by the telcos specifically for carrying voice traffic.

With many voice over IP applications, voice communications will be provided at personal computers and users will tend to use headsets or speakers rather than the traditional handsets, although handsets that can clip onto the video display unit are available. Where loudspeakers are used to give hand free operation, echo control is needed in the computer. This is provided in the latest versions of Microsoft Windows™. In general the use of PCs rather than traditional handsets creates more scope for terminals to be configured incorrectly and correspondingly degrade quality, although it does provide better support for wideband speech.

6.2.1 Speech coding basics

Speech codecs are intended to compress the voice in such a way as to minimally affect voice quality and in turn the quality of service (QoS) delivered by IP telephony systems.

The speech encoder converts the digitized speech signal (after A/D conversion) to a bit-stream, which is packetized and sent over the IP network. The speech decoder then reconstructs the speech signal from the packets received. The reconstructed speech signal is, therefore, an approximation of the original signal. Speech codecs are deployed at end points and so, determine the achievable end-to-end quality.

A speech codec has several important features, including speech quality, bit or compression rate, robustness, delay, sampling frequency, and complexity.

The quality of speech produced by the speech codec will define the upper limit for achievable end-to-end quality. This will determine sound quality for perfect network conditions - no packet loss, delay, jitter, echo, or other quality-degrading factors. Other factors affecting the overall sound quality include the handling of different voices as well as the effect of non-speech signals such as background noise.

Historically, a number of speech codecs have been designed to address BE (Bit Error) occurrence in the communication channel. However, in packet networks the speech codec must be able to deal with lost packets. This ability determines the sound quality in a loaded network and also in congested situations where packet loss is likely to happen.

The delay introduced by the speech coder can be divided into algorithmic and processing delay. The algorithmic delay occurs because of framing for block processing, since the encoder produces a set of bits representing a block of speech samples. Furthermore, many coders using block processing also have a look-ahead function that requires a buffering of future speech samples before a block is encoded. This adds to the algorithmic delay. Processing delay is the time it takes to encode and decode a block of speech samples.

The complexity of a speech-coding algorithm dictates the computational effort required and the memory requirements. Complexity is an important cost factor for implementing a codec and generally increases with decreasing bit rate.

Increasing the sampling frequency from the 8 kHz used for telephony band products to the 16 kHz used for wide-band speech coding produces distinctly more natural, comfortable, and intelligible speech. To date, wide-band speech coding has found limited use in applications such as videoconferencing because speech coders mostly interact with the Public Switched Telephone Network (PSTN) and so have been limited to compatibility with codecs used in the network. There is no such limitation in calls carried wholly over an IP network. Therefore, because of the significant quality improvement attainable, the next generation of speech codecs for VoIP will be wide-band.

6.2.2 Use of traditional circuit switched codecs for voice over IP

The most commonly used codecs for IP telephony today are G.711, G.729A, and G.723.1 (at 6,3 kbps). These were designed for (or based on technology designed for) circuit switched applications.

Mobile telephony has been the major driver for development of speech coding technology in recent years. All the codecs used in mobile telephony, as well as G.729A and G.723.1, are based on the Code Excited Linear Prediction (CELP) paradigm. Due to their design for use in circuit switched networks, these codecs were intended to handle bit errors rather than packet loss.

G.711 was designed for use in circuit switched telephony, and as such it does not include any means to counter packet loss. The common remedy of inserting "zeros" (zero stuffing) whenever packet loss occurs leads to voice break-up and a steep degradation of quality. Error concealment can be introduced by extrapolating/interpolating received speech segments. An example is the method described in Appendix I to ITU-T Recommendation G.711 [5], which provides some improvement but does not guarantee robust operation.

G.729 and G.723.1 belong to the CELP coder class, which is also based on a coding model that was designed for circuit switched mobile telephony. Basic speech quality is worse than PSTN quality.

The CELP coding process uses inter-frame dependencies that lead to inter-packet dependencies. Error propagation resulting from such dependencies leads to poor performance when packets are lost or delayed and so speech quality degrades rapidly with increasing packet loss.

Figure 11 compares the convergence times for resynchronizing encoder and decoder states after packet loss for the traditional G.729 and G.723 codecs and the iLBC codec [5], which has been specified through the IETF AVT (audio/video transport) Working Group. The iLBC codec, which is designed not to have any inter-packet dependencies, recovers much more quickly giving much better performance.

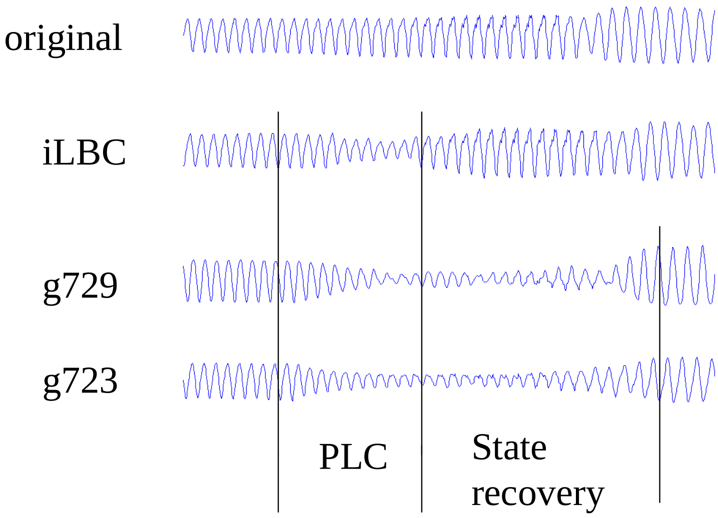


Figure 11: Recovery from packet loss

Subjective evaluations have confirmed that it is more important to restore the state of the decoder after a frame erasure than to attempt to restore the speech that was lost during the frame erasures. Comparisons have also shown that it is always better to use the information in an errored packet than to discard a packet when the error level exceeds a certain threshold.

6.2.3 Speech processing designed for speech over packet networks

6.2.3.1 Introduction

Speech processing software designed for real time communications over IP networks aims to provide an edge-device QoS solution with high voice quality, even under severe network degradations due to jitter and packet loss. Its design is characterized by the following principles:

- speech quality in IP telephony should generally be equal-to or better-than PSTN;
- speech quality should degrade gracefully with increasing packet loss and delay. Moderate packet loss should be inaudible.

6.2.3.2 Codec enhancements

The main approach achieving good speech quality in the presence of packet loss is to avoid inter-packet dependencies and so to prevent error propagation. The new speech processing algorithms support diversity, which results in minimal loss of speech from packet loss but they do so not by adding redundancy as with Forward Error Correction (FEC) methods, but by generating multiple descriptions of the source signal of equal importance, called Multiple Description Coding (MDC) [10]. These descriptions can be decoded independently at the receiver. If all descriptions are received, the source signal can be faithfully reconstructed. If only a subset of the descriptions is received, the quality of the reconstruction is lower

These new techniques achieve a significant increase in performance and can achieve adequate quality at packet loss levels as high as 20 % to 30 %, whereas traditional codecs degrade significantly as packet loss levels rise above 3 % to 5 %.

6.2.3.3 Playout buffer control

Delay variation is the fundamental main problem in real-time voice applications over IP, consequently one of the important functions to be implemented at the receiver is the Playout Controller, which buffers the variable delay (jitter) in the network to give a constant stream of packets for the codec. The playout controller controls the size of the buffer at the receiving end and therefore trades off packet loss and overall delay or latency. The design and algorithms for playout controllers have become much more sophisticated in their ability to minimize additional delay and also conceal packet loss. Furthermore the playout controller can reduce the effect of clock skew, which can occur due to the sender and receiver not being correctly synchronized.

6.2.3.4 "Traditional" Playout Buffer

A traditional playout buffer removes the jitter in the arrival times of the packets by adding delay so that the total delay in the network and the buffer is constant. The objective of a playout buffer algorithm is hence to keep the buffering delay as short as possible while minimizing the number of packets that arrive too late to be used. A larger playout buffer causes increase in the delay and decreases the packet loss. A smaller playout buffer decreases the delay but increases the packet loss.

The traditional approach is to store the incoming packets in a buffer (packet buffer) before sending them to the decoder. The most straightforward approach is to have a buffer with a fixed number of packets. This results in a constant system delay (if there is no clock drift) and requires no computations and therefore gives a minimal complexity. The drawback with this approach is that the length of the buffer has to be made sufficiently large to accommodate the maximum jitter (which in practice occur quite seldom).

In order to keep the delay as short as possible it is important that the jitter-buffer algorithm adapts rapidly to changing network conditions. Therefore, playout buffers with dynamic size allocation, so called adaptive playout buffers, are most common nowadays.

The adjustment of delay is achieved by inserting packets in the buffer, when the delay needs to be increased, and removing packets when the delay can be decreased. The insertion of packets usually consists of repeating the previous packet. Unfortunately, this will almost always result in audible distortion and hence most adaptive playout buffer algorithms are very cautious when it comes to delay adaptations in order to avoid such effects. To avoid audible distortion, the removal packets can only be done during periods of silence. Hence, delay builds up during a period of speech and it can take several seconds before a reduction in the delay can be achieved. Also, high delay at the end of a period of speech will have a severe effect on the conversation since it increases the probability of double talk. Conversational tasks would therefore become much easier if the delay could be adapted during active speech and be kept to a minimum at the end of each period of speech.

This traditional packet buffer approach is limited in its adaptation granularity by the packet size since it can only change the buffer length by adding or discarding integral numbers of packets.

Some of the current implementations of adaptive playout buffers have been shown to experience problems when there are packet losses in the network. For example, studies for TIPHON in [14] show that the playout buffer delay can increase a lot for cases where packet losses are present.

6.2.3.5 Packet Loss Concealment

Until recently, two simple (codec independent) approaches for packet loss concealment have been used.

The first method, referred to as Zero Stuffing (ZS), is obtained by simply replacing a lost packet with a period of silence of the same duration as the lost packet.

The second method, referred to as packet repetition (PR), assumes that the difference between two consecutive speech frames is quite small. Hence, the lost packet is replaced by simply repeating the previous packet. In practice, though, even a minor change in, for example, the pitch frequency is easily detected by the human ear. In addition, it is virtually impossible to achieve smooth transitions between the packets with this approach. However, this approach performs fairly well for very small probabilities (less than 3 %) of packet loss. Packet repetition outperforms zero stuffing but both methods are very sensitive to packet loss compared to the more advanced methods.

Recently, the ITU standardized a method for packet loss concealment in G.711 Appendix I (usually referred to as G.711 PLC). This is a more sophisticated method that tries to estimate the lost packet from previously decoded speech and hence cannot be implemented in the packet buffer.

Some codecs, such as those based on CELP, have their own built-in packet loss concealment algorithm. In many cases this gives a reasonable concealment during the loss. Unfortunately, as previously noted, many of these codecs suffer from their packet inter-dependencies instead.

6.2.3.6 Clock drift (skew)

Clock drift is the difference in the rates of the clocks at the sending and receiving end. The traditional approach (in a TDM network) is to deploy a clock synchronization mechanism at the receiver to correct for clock drift by comparing the number of samples received with the local clock. In an IP network, however, it is hard to do reliable clock drift estimation. The reason is that the estimates of drift only can be based on averaging packet arrivals at a rate of typically 30 to 50 per second instead of averaging on a per sample basis at a rate of 8 000 per second (as done in TDM networks). In addition, because of the jitter present in IP networks it is almost impossible to obtain an accurate estimate of the clock drift and hence many algorithms designed to mitigate this effect fail.

6.2.3.7 Advanced Algorithms

Advanced algorithms are coming onto the market that include both delay adaptation and error concealment in one unit and adapts quickly to changing network conditions to ensure high speech quality with minimal buffer latency. An example is NetEQ (Network Equalizer) (Ref:http://portal.etsi.org/stq/presentations2003/06HenrikAstromPresentation%20.pdf). The following is a generic discussion of the techniques that can be used. These algorithms work on both the input and the output of the decoder, whereas traditional algorithms worked only on the inputs. By working on the outputs, these algorithms are able to smooth the speech when a packet is lost rather than just replacing the lost packet.

Figure 12 shows a simple block diagram of an advanced algorithm and contrasts it with traditional packet loss concealment.

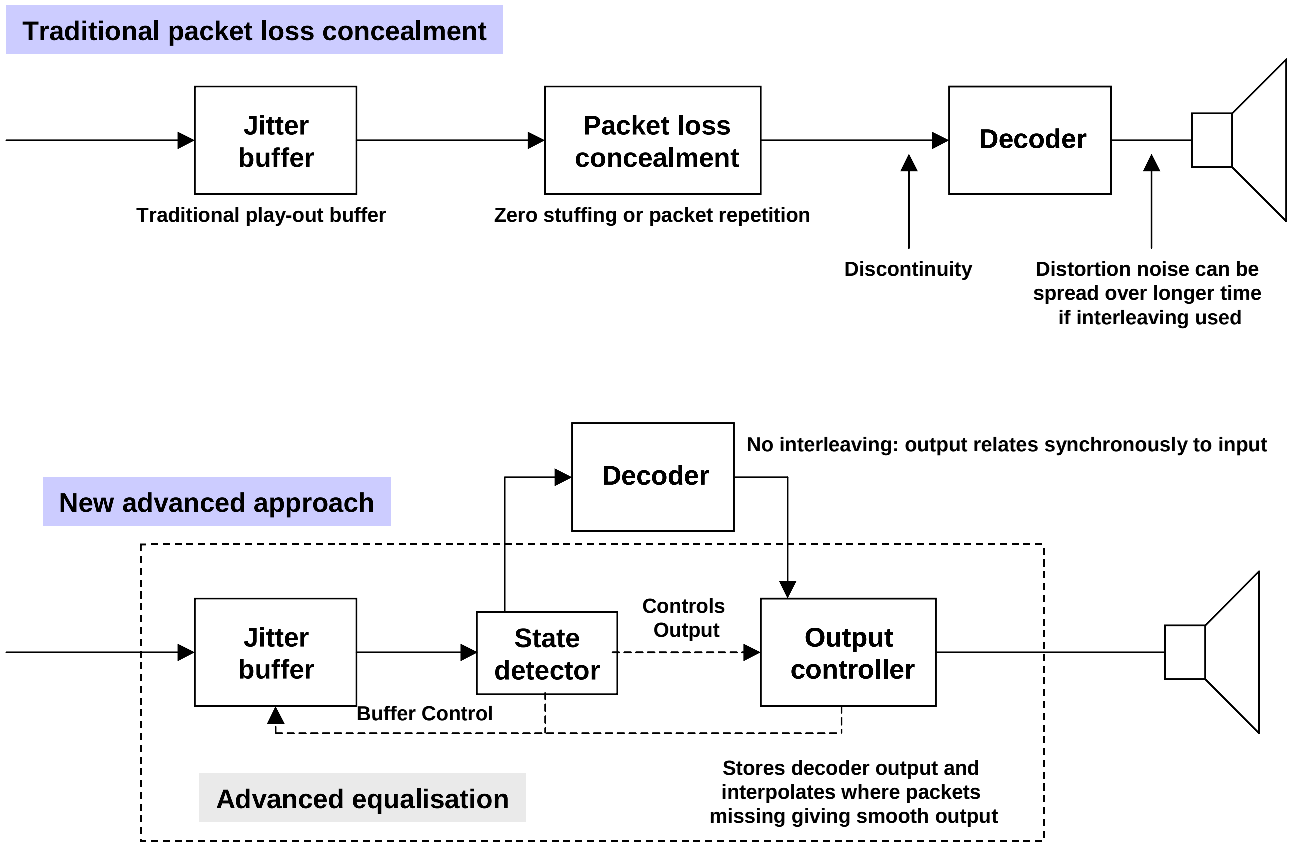


Figure 12: Simple block diagram depicting how an advanced algorithm can interacts with the speech decoder and the jitter buffer

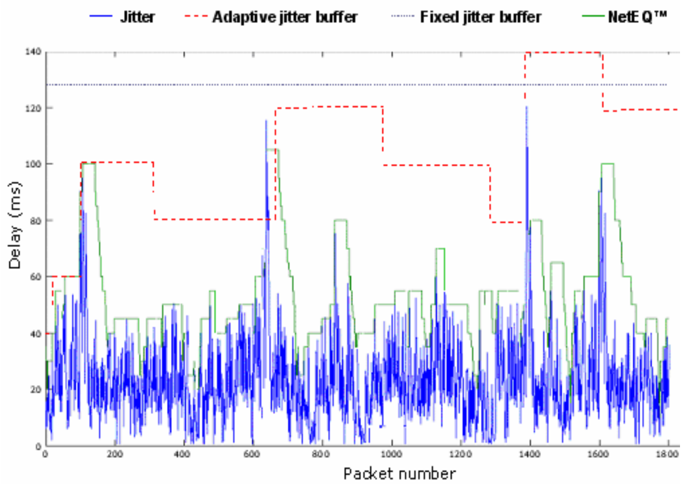
The algorithm works in the following way. When packets are available from the jitter buffer, the packets are sent to the decoder and the decoder output is used but with some additional fixed delay in the output controller. When packets are not available (because the jitter buffer has emptied as a result of increasing network delay) the packet loss detector stops the output controller from using the output of the decoder and instead the output controller inserts a new short synthetic speech segment instead of the output of the decoder to interpolate across the gap caused by the packet loss. The purpose of the additional delay added in the packet controller is to allow for this interpolation.

The effectiveness of this approach depends on the coder not using spreading so that the effects of lost packets on the decoder output are not spread over a significantly greater time period than the duration of the lost packets. The main advantage over traditional systems is that the concealment takes place on the decoder output rather than the input and includes interpolation rather than extrapolation allowing for the output to be smoothed.

The loss detector and the output controller are able to measure the extent and jitter at the network output and they are used to control the jitter buffer so that the size of the buffer can be reduced when jitter is low. Changes to the jitter buffer require corresponding changes to the speech output as described in clause 6.2.3.4.

The superior performance in terms of clock drift is achieved because the algorithm does not have to estimate the actual clock drift but is able to mitigate its effect automatically at the speech output rather than the decode input. It is therefore a very attractive alternative to a standard clock synchronization mechanism.

Studies were carried out to compare the delay of the NetEQ algorithm with alternative solutions. The graph below illustrates the delay performance of different algorithms on a channel with quite a lot of jitter. NetEQ manages to keep the delay much lower than both the adaptive and the fixed playout buffer. In general a delay improvement of 30 ms to 80 ms can be expected with the NetEQ algorithm compared to traditional approaches. Figure 13 also shows that NetEQ adapts to the envelope of the jitter very efficiently, since the delay is reduced in less then a second after a jitter peak.



NOTE: That the constant system delays of about 80 ms have been removed from all curves.

Figure 13: Delay performance for different playout buffers on a system with 20 ms packets

6.2.3.8 Other approaches

Other approaches to deal with packet loss are forward error correction schemes and interleaving schemes.

Forward error correction schemes are designed to correct bit errors that are well distributed in time, whereas packet loss results in errors in many consecutive bits (a burst), which significantly decreases the efficiency of FEC schemes. In order to combat a burst or errors, redundant information has to be added and spread over several packets, which introduces greatly increased delay. Hence, the repair capability of forward error correction is limited by the delay budget. Furthermore the use of additional bits increases the network loading and so may aggravate the problem whose effects the scheme is trying to reduce.

Another technique for reducing the effect of packet loss is interleaving where the coder output frames are interleaved to that a burst of errors is spread over more than one frame. Whilst interleaving does not increase the data rate of transmission it does increase delay significantly. The efficiency of loss recovery increases if the source packet is interleaved and spread over more packets but the more packets that are used the higher the delay.

Both forward error correction and interleaving are used in GSM.

7 Network design

7.1 Introduction

The approach to network design depends on the objective that the designer is trying to achieve. The common objective is to improve the quality for voice transmission, but telcos in particular are interested in being able to guarantee that a specified level of performance is achievable. Where they can guarantee a level of quality, some may be interested in offering different levels of quality at different tariffs.

In circuit switched networks, the transmission quality was normally constant and independent of traffic once a call was established. Traffic congestion meant that set-up might be blocked. Packet technology offers the possibility to trade quality against capacity and so when there is a high level of congestion there is the option to have communications of reduced quality or to wait until the congestion eases and then communicate at better quality. There is little knowledge of how the various choices may be presented to users and how the users will react to these options, as the choices have not been available with circuit switched technology. Intuitively the response is likely to depend on the circumstances of the caller and the party called. Callers with an urgent need to communicate will accept any quality that is intelligible whereas users whose communications are not urgent may prefer to wait for better quality especially if the main purpose is the pleasure of talking rather than a more functional requirement.

Network design, measurement and control issues can be handled at two levels:

- the service or application level;
- the transport or network level.

The work in ETSI has focussed on QoS signalling as part of call set-up at the application level, whereas work in IETF is almost exclusively at the network level.

It is important to distinguish between the telco model and the Internet model when considering network design. Techniques that apply to the application level or that make a network "application conscious" are relevant only where the telco model is being followed. Such techniques are likely to make interconnection more complex and this is a cost that needs to be taken into account in the overall design.

A further consideration at the service or application level is the substantial influence of the terminal design and especially the extent to which codecs can tolerate variable delay and packet loss in the networks. This means that the benefits of schemes that relate only to public networks and do not include the end networks and the terminals may be quite limited.

7.2 Guarantees

The word "guarantee" is used with two quite different meanings:

- an "absolute" guarantee;
- a "statistical" guarantee.

In practice since the traffic demand is statistical and in effect unbounded, there can be no absolute guarantee of both quality and availability. Therefore an absolute guarantee means in practice that access may have to be denied to others and even also to those using the guaranteed service.

A statistical guarantee means that the network will be designed to achieve a certain quality level for x % of the time assuming a given level of demand.

With packet networks call quality may change during the course of a call, whereas with circuit switched technology it remains constant. With an absolute guarantee, the quality should not fall below the guaranteed level since the network loading will be controlled by denying access to those calls that would cause congestion.

The provision of an absolute quality guarantee requires call related QoS signalling at the application level.

7.3 Call related Quality of Service signalling

There are two approaches to call-related QoS signalling:

- negotiating QoS during call set-up with the possibility of failing the call set-up attempt if the QoS is not adequate;
- passing forward QoS information to guide routing or other decisions in networks that handle the call subsequently.

7.3.1 Basic concept for call related signalling negotiation

The main work on QoS signalling negotiation has taken place within the ETSI TIPHON project.

NOTE: It finished before TISPAN started.

The following gives a brief description of the main concepts. Figure 14 shows an example of how the signalling would work.

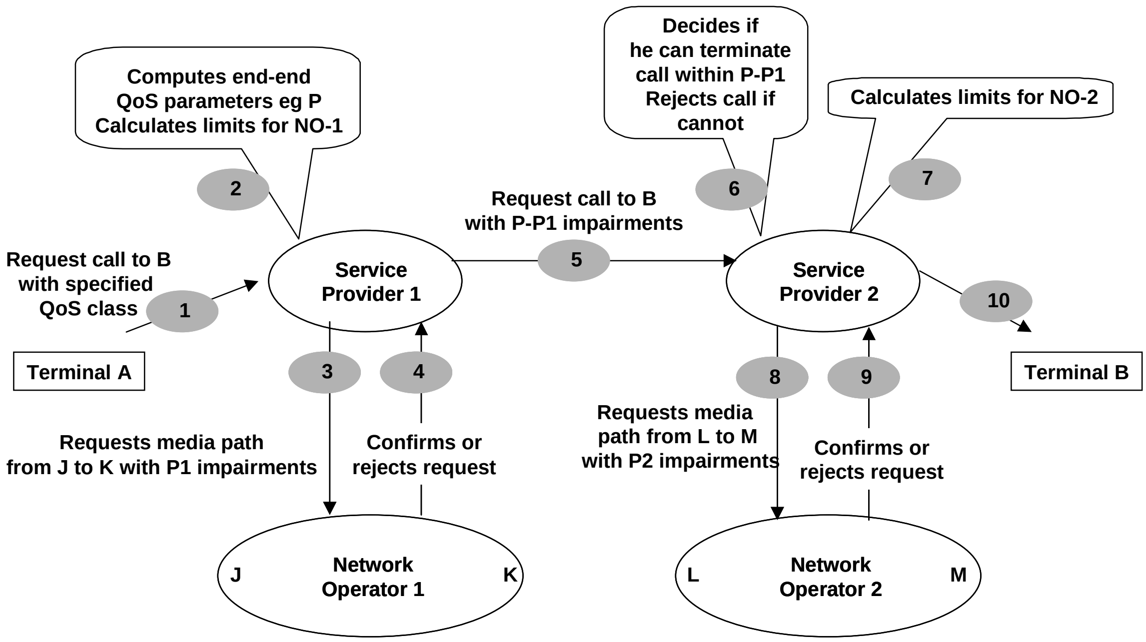


Figure 14: Operation of QOS signalling

The caller selects the desired QoS class and the main service provider transforms this class into transport impairments values. The service provider then starts a process of negotiation with network operators and other interconnected service providers.

The call/call QoS signalling between service providers is not a dialogue. In the forward going SETUP message, the sending service provider specifies the values of the impairment parameters within which it wishes the receiving service provider to complete the call (i.e. The difference between The end-end impairment parameter limit (e.g. 200 ms) and the accumulated value for the impairment parameter)

The receiving service provider then decides whether or not to continue the set-up procedure and route the call onwards, or whether to fail the call on the basis of inability to meet the QoS objectives. This decision may use stored information on the recent performance, e.g. which destinations can be reached with what impairments.

If the call is failed, then the RELEASE message passes back to the sending service provider. The sending service provider may decide to pass the release back further or to recommence set-up with another service provider.

Thus the QoS signalling does not involve any additional message flows, just the inclusion of specific information in the messages that have already been defined for call control.

7.3.2 Discussion of the usefulness of call related QoS signalling negotiation

The precise role of QoS signalling is not fully stable and there are slightly different ways in which it could be used.

The main function of QoS signalling is to implement the absolute quality guarantee and to fail call set-up where the guaranteed quality cannot be met. It also could be used to enable the service provider to adjust the charging where the guaranteed quality is not met. Different views are held as to whether this is helpful to the caller. One school of thought says that it is always better to connect the call and let the user assess whether he wants to continue with the call since this policy gives the user the maximum choice and lets him take into account the urgency of the communications.

There has also been discussion in TIPHON on the effectiveness of QoS signalling in general purpose networks where delay sensitive real-time speech can be given priority in the router queues. Assume first that less than say 20 % of the traffic is delay sensitive. This means that prioritization should have a substantial effect in reducing variable delay. The issue is then whether the potential QoS problem is caused by:

- the fixed delay that is a consequence of the network design, topology and the physical distances involved; or
- the variable delay caused by the queues in the routers.

If the main problem is the fixed delay, then the networks may be fundamentally unsuitable for the traffic being considered. This could be determined a priori from knowledge of the design and does not require call-related signalling, although call related signalling could help with route selection.

If the main problem is the variable delay, then prioritization may solve the problem in which case QoS signalling does not add much real value. If the problem cannot be solved, then QoS signalling may help by screening the calls and denying access when the networks are too congested, but fundamentally the networks concerned cannot be relied on as a suitable means of carrying the traffic.

These arguments are summarized in the flow chart in figure 15.

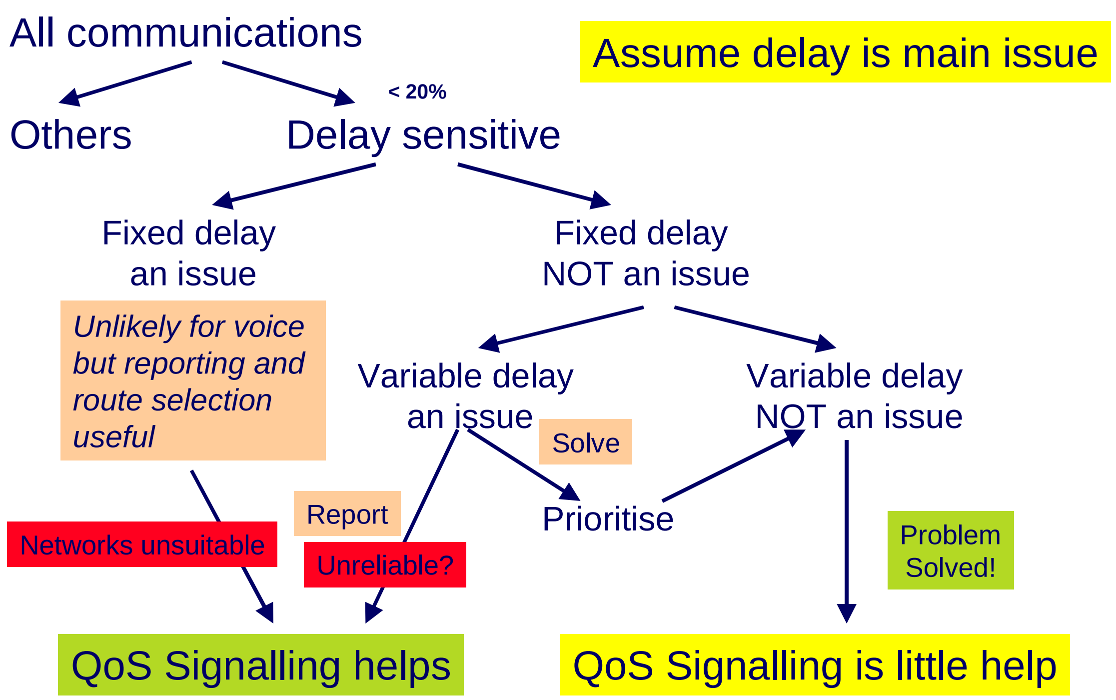


Figure 15: Effectiveness of call related QoS signalling

A further important point is that call related QoS signalling needs to be supported end-to-end to be of any use. Since calls are likely to traverse more than one network and many calls will traverse at least 5 networks, two end networks, two access networks and one transit network, the introduction of QoS signalling faces the problem of achieving critical mass i.e. of becoming sufficiently widely supported that there is a high probability of QoS signalling being available end-to-end.

7.3.3 Conclusion on call-related QoS signalling negotiation

Call-related QoS signalling is not a mature concept and there is no practical experience of the extent to which users would value the functionality that it provides. QoS signalling is designed to support the concept of a service with an absolutely guaranteed QoS, but this means access may be denied at times. It is not at all clear that users will prefer this approach to one where they always have service access but quality may vary.

The analysis of the quality factors suggests that QoS signalling is only likely to be useful if the network performance is marginal. Provided that the proportion of delay sensitive traffic to the total traffic is low, the use of prioritization should be effective in ensuring adequate performance for delay sensitive traffic in networks that are designed to be fundamentally suitable for such traffic, i.e. where the fixed delays are within an acceptable level.

A further point is that the concept depends on an knowledge of the relationship between the terminals and the network performance and this relationship may not be known at all of only known very approximately by the entities making the control decisions.

Thus in summary the basic concept is open to considerable doubt from the perspective of the user and the proportion of cases where it might add value is likely to be low. Thus overall it does not appear to be worth pursuing.

7.4 Signalling for congestion control

The alternative to call related signalling for quality control purposes is management signalling to signal congestion and to tell services what are originating new calls to block them until the congestion reduces. This feature is considered to be of high priority for the control of congestion arsing from events such as mass call-in programmes where traffic peaks of some 20 times the normal levels can be experienced. The objective is to block the call attempts as close to the origin of the traffic as possible and so to reduce the ladings on the network. The main concern is that the effective capacity of the network can fall as it goes into overload conditions creating an unstable "slippery slope" situation. There is a risk that if processors become overload they fail uncontrollably and have to be restarted creating a prolonged outage.

The control of congestion and mass calling events is clearly important, but it is questionable if it should be treated in the same way as the control of network performance.

7.5 Reservation, Segregation and prioritization of traffic type

The sources of the impairments generated in the networks are the queues that delay the packets at the routers. Queues may be a problem in networks of all sizes, but at points where the transmission capacity is low they are especially critical because delay sensitive packets may be held up for significant lengths of time while large data packets are transmitted. These problems can occur especially in access networks, and some access systems will use techniques for subdividing long data packets to prevent long delays to smaller speech packets.

Three quite different solutions have been developed in IETF:

- RSVP, which reserves capacity on a per-call basis;
- DiffServ, which applies different behaviours at the routers to different classes of traffic;
- MPLS (Multi-Protocol Labelling System), which segregates traffic by adding labels to packets so that the internal routers can route on the labels. This has the effect of segregating a network into several separate virtual networks that can each be dimensioned differently to that for example the virtual network for speech can be dimensioned generously for low router queues.

7.5.1 Reservation (RSVP)

RSVP is the only reservation protocol under consideration. It is relatively complex to manage because it is call related and does not scale well and is therefore only being considered for use in access networks where bandwidth may be low.

7.5.2 Prioritization (DiffServ)

DiffServ is simple prioritization concept where different classes of traffic are assigned to different behaviours in terms of their handling in the switches. Initially DiffServ was developed only for use internally within a network and so the classes and behaviours had only local meaning, but more recent work in IETF is standardizing some classes. The effectiveness of DiffServ will depend very much on the types of traffic carried in the network. If the majority of the traffic is not delay sensitive, then DiffServ could be very effective in improving quality. If most or all of the traffic is delay sensitive, for example because the non-delay sensitive traffic is being handled separately over the Internet, then prioritization will not improve quality for speech traffic but could be useful for calls to emergency services. It is therefore extremely important for the development of appropriate network performance that the relationship between the NGN and the Internet is clarified, e.g. will the NGN carry non delay-sensitive traffic?

7.5.3 Segregation

It is not easy to see how segregation will help network performance as each the network will need to be dimensioned separately for each logically segregated traffic type and will need to take account if the statistical variations of the traffic. Statistically it is normally better to aggregate traffic of different types because according to Erlang's work the additional capacity needed to handle traffic peaks with the same probabilities reduces as the total level of traffic increases. It is not clear whether MPLS can be used for a dynamic rather than a fixed method of segregation, but this starts to equate to prioritization.

8 Network performance targets

The main standard that defines network performance for international connections from UNI to UNI is ITU-T Recommendation Y.1541 [15]. This Recommendation defines six different network QoS classes that are unrelated to the TIPHON classes. For each class the Recommendation specifies performance values for each of the IP-related performance parameters defined in ITU-T Recommendation Y.1540 [14]. The network QoS classes defined here are intended to be the basis of agreements between end-users and network service providers, and between service providers. The limited number of QoS classes defined in Y.1541 support a wide range of applications, including the following: real time telephony, multimedia conferencing, and interactive data transfer.

Y.1541 defines limits for the following network performance parameters:

- IPTD - IP packet transfer delay;
- IPDV - IP packet delay variation;
- IPLR - IP packet loss ratio;
- IPER - IP packet error ratio.

The IP network QoS classes that have been defined provisionally based on these parameters are shown in table 1.

Table 1: IP network QoS classes in Y.1541

Network Performance Parameter	Nature of Network Performance Objective	QoS Classes: Class 0	Class 1	Class 2	Class 3	Class 4	Class 5 Un-specified
IPTD	Upper bound on the mean IPTD	100 ms	400 ms	100 ms	400 ms	1 s	U
IPDV	Upper bound on the 1-10^{-3} quantile of IPTD minus the minimum IPTD	50 ms	50 ms	U	U	U	U
IPLR	Upper bound on the packet loss probability	1×10^{-3}	1×10^{-3}	1×10^{-3}	1×10^{-3}	1×10^{-3}	U
IPER	Upper bound	1×10^{-4}					U

General notes:

The objectives apply to public IP Networks. The objectives are believed to be achievable on common IP network implementations. The network providers' commitment to the user is to attempt to deliver packets in a way that achieves each of the applicable objectives. The vast majority of IP paths advertizing conformance with ITU-T Recommendation Y.1541 [15] should meet those objectives. For some parameters, performance on shorter and/or less complex paths may be significantly better.

An evaluation interval of 1 minute is provisionally suggested for IPTD, IPDV, and IPLR, and in all cases the interval must be reported.

Individual network providers may choose to offer performance commitments better than these objectives.

"U" means "unspecified" or "unbounded". When the performance relative to a particular parameter is identified as being "U" the ITU-T establishes no objective for this parameter and any default Y.1541 objective can be ignored. When the objective for a parameter is set to "U", performance with respect to that parameter may, at times, be arbitrarily poor.

Table 2 gives guidance for the applicability and engineering of the network QoS classes in order to support different applications.

Table 2: QoS classes in Y.1541

QoS Class	Applications (Examples)	Node Mechanisms	Network Techniques
0	Real-Time, Jitter sensitive, high interaction(VoIP, VTC)	Separate Queue with preferential servicing, Traffic grooming	Constrained Routing and Distance
1	Real-Time, Jitter sensitive, interactive (VoIP, VTC)		Less constrained Routing and Distances
2	Transaction Data, Highly Interactive, (Signalling)	Separate Queue, Drop priority	Constrained Routing and Distance
3	Transaction Data, Interactive		Less constrained Routing and Distances
4	Low Loss Only (Short Transactions, Bulk Data, Video Streaming)	Long Queue, Drop priority	Any route/path
5	Traditional Applications of Default IP Networks	Separate Queue (lowest priority)	Any route/path

The QoS classes in Y.1541 are not levels of quality like the TIPHON classes, but rather QoS requirements profiles for different types of traffic. The general quality level of the requirements is high. For example the packet loss levels are very low, and for speech would allow use of codecs such as G.711 that are designed for circuit switched applications.

One concern is that the work to date on network performance has not specified levels of jitter, yet jitter and its spectrum is the most critical parameter in terms of its influence on performance and the distortion perceived by the user. The dominant interaction between the terminal and the network is the way in which the terminal handles jitter. The problem is that jitter is the least understood parameter.

9 End-to-end QoS classes at the application level

The main work on end-to-end QoS classes has been carried out in the ETSI TIPHON project and is described in TS 101 329-2 [1]. Three classes of end-to-end speech QoS are defined for TIPHON systems: WIDEBAND, NARROWBAND and BEST EFFORT. The TIPHON speech QoS classes WIDEBAND and NARROWBAND will provide performance guarantees for 95 % of all connections (i.e. a statistical guarantee). The BEST EFFORT class provides no speech performance guarantees. The classes are defined from mouth-to-ear and therefore include the transit and access networks, the end networks and the TIPHON terminal characteristics. Each of the classes defined is specified by three performance metrics: Overall Transmission Quality Rating (R), Listener Speech Quality (One-way non-interactive end-to-end Speech Quality) and End-to-end (mean one-way) Delay. Table 3 shows the classes.

Table 3: TIPHON QoS classes

	Wideband	Narrowband			Best Effort
		High	Medium	Acceptable	
Description	IP telephony service using wideband codecs (codecs encoding analogue signals with bandwidth in excess of 3,1 kHz) and QoS-engineered IP networks	IP telephony service provided via QoS–engineered IP networks			Voice communication service operated over non QoS-engineered IP networks such as the public Internet
	Supposed to be better than the PSTN	***Supposed to be equivalent to recent ISDN services***	Supposed to be equivalent to recent wireless mobile telephony services in good radio conditions (EFR codec)	Supposed to be equivalent to common wireless mobile telephony services (FR codec)	Supposed to provide usable communications service but will not provide guarantees of performance (with periods of significantly impaired speech quality, and large end-to-end delays which are likely to impact the overall conversational interactivity)
Overall Transmission Quality Rating (R)	**Not applicable**	**> 80**	**> 70**	**> 50**	**> 50** Target value
Listener Speech Quality	**Better than G.711**	**Equivalent or better than ITU-T Recommendation G.726 [7] at 32 kbit/s**	**Equivalent or better than GSM-FR**	**Not defined**	**Not defined**
End-to-end Delay	**< 100 ms**	**< 100 ms**	**< 150 ms**	**< 400 ms**	**< 100 ms** Target value

The TIPHON classes are unusual in that they involve a combination of two independent parameters (delay and listener speech quality) and the overall transmission quality rating, which includes the delay and listener speech quality parameters as well as other parameters. Thus the three parameters are not orthogonal as might be expected.

From a design perspective, a designer has some control over the delay by choosing the codec algorithm including the jitter buffer play-out, the routings and network loadings, and over listener speech quality (distortion) by choosing the codec, bit rate and network loadings. He can thus adapt his design to achieve a particular overall quality level.

Figure 16 shows the quality classes and the inter-relationship of the parameters where:

- Purple = Best effort;
- Red = Acceptable;
- Yellow = Medium;
- Green = High.

In order to draw figure 16, typical values of other parameters have been used when calculating the OVR with the E-Model.

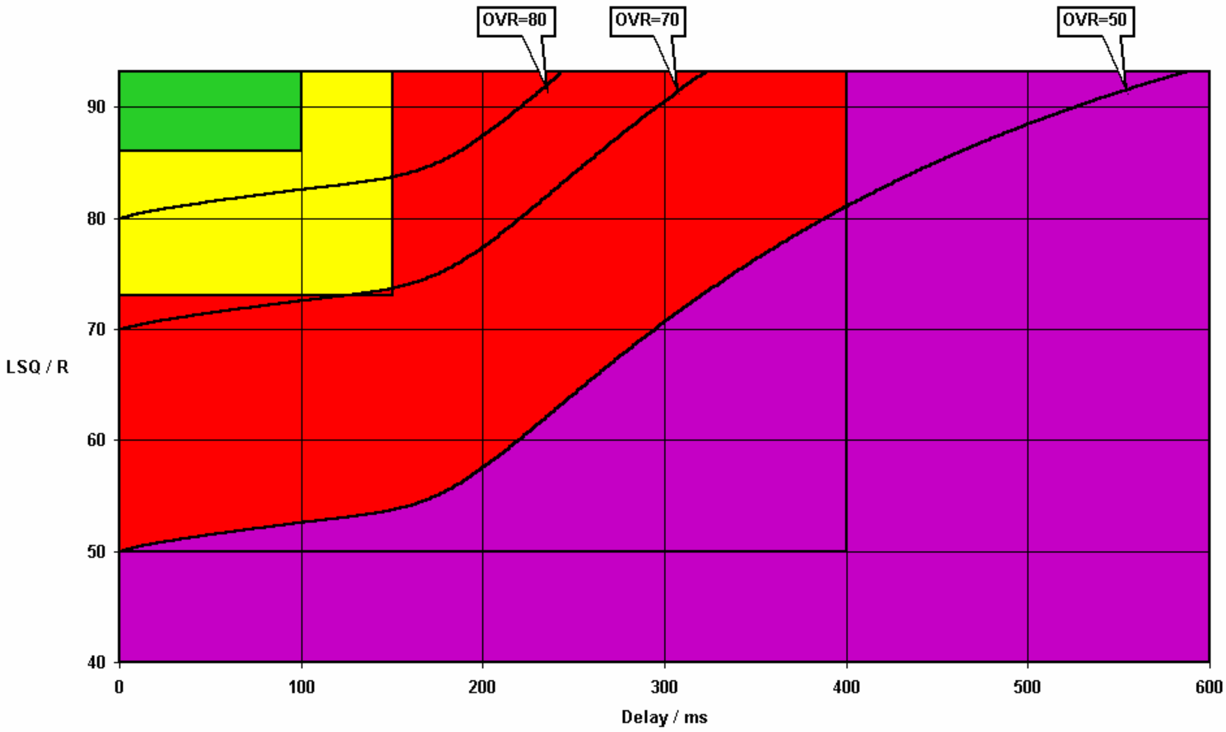


Figure 16: TIPHON QoS classes

In practice the figure shows that for the medium and high classes (yellow and green) the constraint on the OVR has little or no effect since the requirements on delay and LSQ alone define the classes. The understanding that has developed of user views is that the users tend not to distinguish individual parameters when assessing a connection and that it would be better to use Overall Transmission Rating alone when defining classes and thus the classes shown in figure 17 would be more appropriate.

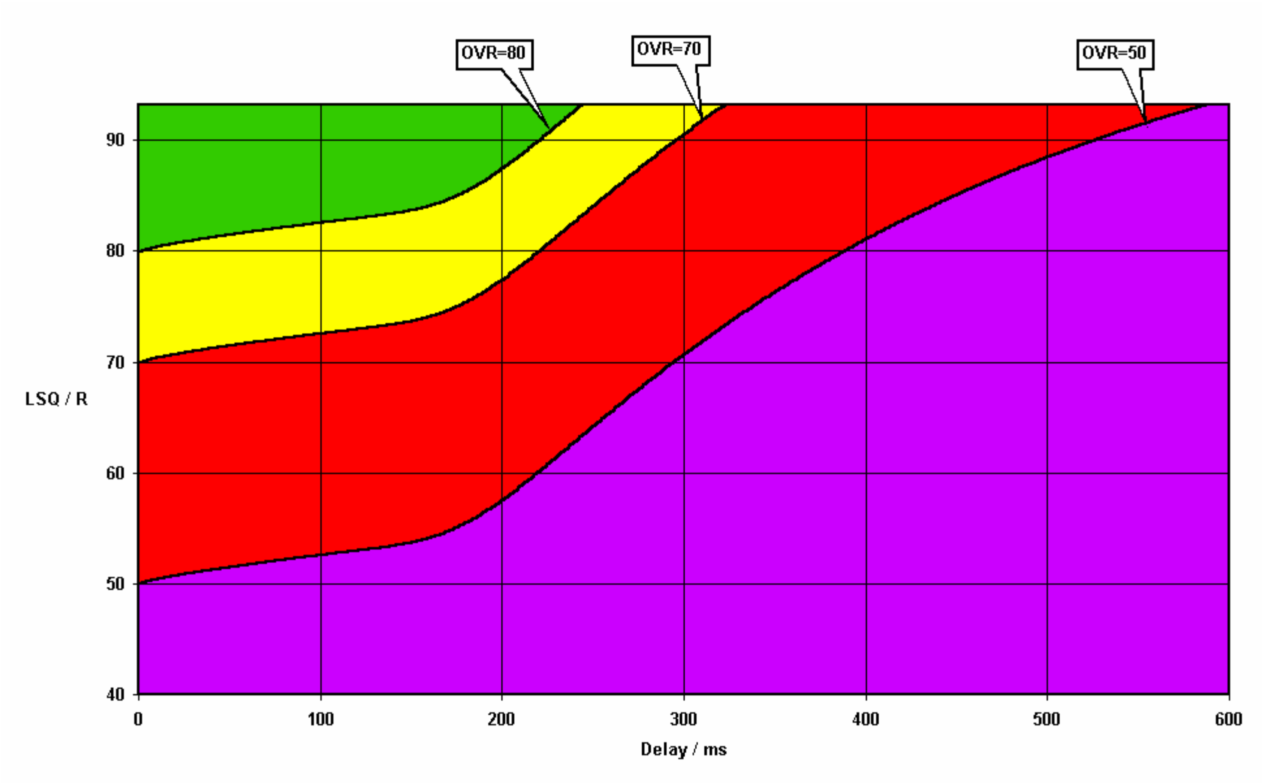


Figure 17: Classes based on Overall Transmission Ratings alone

TIPHON has also produced a design guide: TR 102 024-7 V2.1.2. In this document transmission planning and network maintenance principles are listed as well as guidance on main transmission parameters.

Although the TISPAN has superseded TIPHON, TISPAN has not revised this work.

10 Testing

A range of different approaches should be considered for testing:

- acoustical interface to acoustical interface, i.e. Mouth-to-Ear;
- acoustical interface to electrical interface, i.e. Mouth-to-digital network interface;
- electrical interface to acoustical interface, i.e. digital network interface-to-Ear;
- "Half channel" approach for accessing the packet interface (send or receive), see clause 10.1.

10.1 "Half channel" approach

The half channel approach describes a testing methodology accessing the IP packet stream in order to be able to define transmission requirements between the IP packet stream and the acoustical or electrical interfaces of IP endpoints, e.g., VoIP phones or VoIP gateways.

Transmission testing of VoIP devices, done as a system test, only; i.e. tests are carried out between the electrical interface of the gateway and the acoustical interface of the VoIP phone or between the acoustical interfaces of two VoIP phones, does not disclose the required detail of information on transmission impairments and their possible causes.

For a proper determination of the amount of speech quality degradation caused by a single VoIP endpoint, half channel measurements are necessary.

This could be achieved by defining a half channel measurement adaptor, which, in receive direction, is collection the RTP data via a packet filter and which, in send direction, is inserting new packets into the RTP stream. The RTP data stream has to be encoded and decoded under defined real time conditions; i.e. the delay incurred by the half channel measurement adaptor has to be known and must be kept constant. See figure 18.

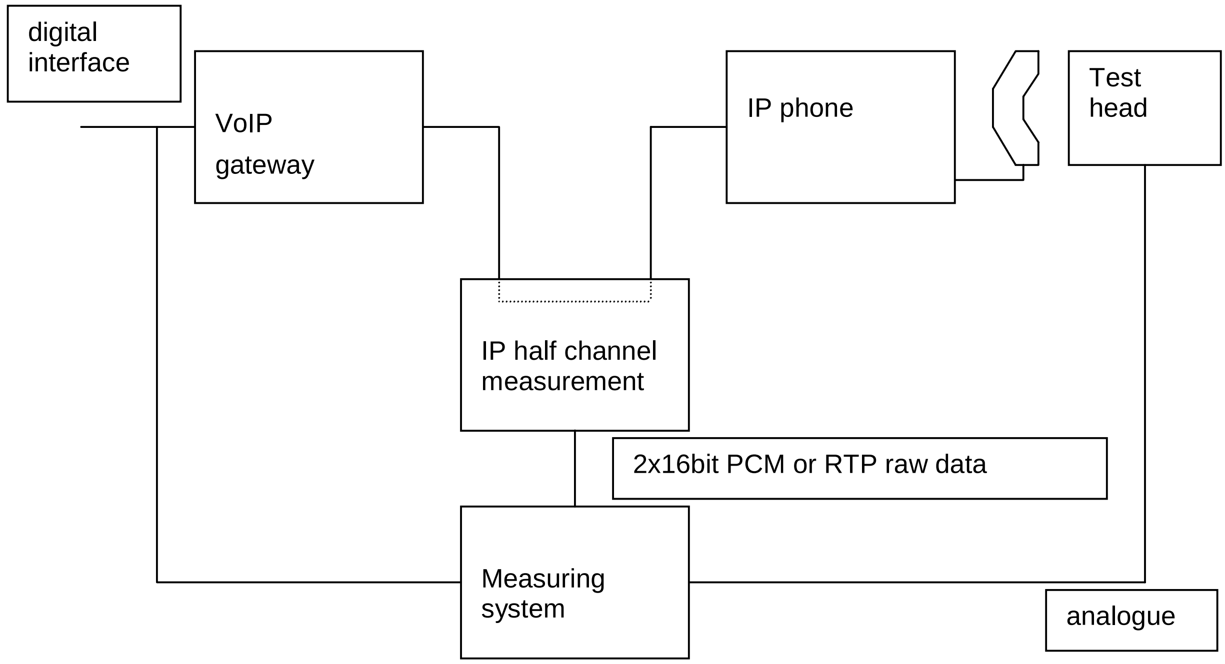


Figure 18: Test configuration

Figure 18 shows the application of the half channel measurement adaptor. The "Measuring system" device can perform measurements for the VoIP phone or for the gateway, separately.

10.2 Quick and Indicative Testing Suite

10.2.1 General Test Description

Measurements should be done with three basic configurations:

- Electrical - Electrical Connection;
- Acoustical - Electrical Connection;
- Acoustical - Acoustical Connection.

Measurements should be conducted using two kinds of input signals:

- real speech samples;
- artificial test signals (according to ITU-T Recommendation P.501 [12]).

For all conditions, qualified speech material (at least two different languages, sample size of about 30 seconds) should be used to achieve a good comparability between different test conditions. All speech files of an test suite should be evaluated and archived in order to allow listening and further signal analysis.

In order to reproduce realistic conditions for acoustical end to end quality measurements both subscribers should be substituted by dummy heads (Head And Torso Simulators, HATS) during the tests, each equipped with an artificial mouth and artificial ears (type 3.4). The positioning of handsets should follow ITU-T Recommendation P.64 [11]. Measurements using the electrical interfaces should be carried out in the same way. Test signals and methods are specially developed to determine ***instrumental quality parameters influencing the conversational quality*** like double talk performance, switching characteristics, echo performance and others.

10.2.2 Tests Based on Instrumental Assessment of Speech Samples

For speech samples that are recorded according to the measurement scenarios described in this clause, instrumental speech quality evaluations should performed. Latest psychoacoustic instrumental analyses of electrical-electrical VoIP transmission scenarios using the Telecommunications Objective Speech Quality Assessment method TOSQA as well as ITU-T Recommendation P.862 [13] lead to an one dimensional test results with a high correlation to auditory perceived speech sound quality. These methods have been validated for VoIP transmission scenarios and are therefore applicable for those scenarios using electrical interfaces. TOSQA has been used and validated in the first and second ETSI VoIP Quality Test Events. Here it demonstrated the same accuracy compared with P.862.

For recordings at the acoustical interface, i.e. for electrical-acoustical as well as for acoustical-acoustical VoIP transmission scenarios, TOSQA (or in future possibly ITU-T Rec. P.AAM) should be used, because P.862 has not been verified for these scenarios.

10.2.3 Tests Based on Auditory Assessment of Speech Samples

The auditory test results lead to speech quality ratings expressed through mean opinion scores MOS-LQS. These MOS-LQS values represent the average test result derived from all individual ratings of a group of untrained test persons, who assess the auditory perceived speech sound quality. Auditory assessments should be performed with a subset of the recorded speech material in order to confirm the instrumental test results.

10.2.4 Tests Based on Instrumental Computational Assessment Using Speech like (P.501) Test Signals

The auditory perceived quality for speech controlled, non-linear or time-variant systems is influenced by additional parameters like echo disturbances, double talk performance, switching characteristics, background noise transmission and others. These parameters like talking-related impairments (e.g. echo) or conversational aspects (e.g. double talk performance) determine the overall quality of the complete system. Tests based on sophisticated test signals and analysis methods were developed to determine the corresponding instrumental parameters. Depending on the interfaces used during the tests (acoustical, electrical) parameters according to the following list should be measured:

- one-way delay in send and receive direction;
- send loudness rating SLR, receive loudness rating RLR, junction loudness rating JLR, overall loudness rating OLR ("mouth to ear");
- frequency responses and distortion, switching characteristics like minimum activation level, sensitivity of double talk detection;
- double talk performance;
- background noise transmission at idle mode, with near end signal, with far end signal;
- echo delay, single talk echo, double talk echo;
- evaluation of the quality of implemented PLC and jitter buffer design using cross-correlation analysis and the "Relative Approach". The cross-correlation analysis is suited to analyze and demonstrate the current implementation whereas the Relative Approach is a hearing model based method to determine audible disturbances introduced by PLC or the jitter buffer control in the time and frequency domain. Consequently the combination of both methods can be used to optimize the current implementation.

These tests should be performed in order to determine instrumental quality parameters for a given connection. The instrumental tests for the determination of implemented parameters is meant to check common requirements in telephonometry and to identify parameters which may lead to auditory perceived conversational quality degradation.

Listening examples should be generated to document the findings during the tests.

The results should be compared to requirements recommended in the relevant standards (ETSI, ITU-T).

10.2.5 Specific Echo Measurements

Detailed tests of echo cancellers and performance tests for background noise transmission will should be carried out.

For the echo cancellers implemented in gateways a hybrid providing different ERL values (echo return loss) should be connected. Tests should be carried out under the following conditions:

- variation of ERL between 6 dB and 40 dB;
- the most important echo canceller test according to G.168 (convergence characteristics, convergence with low level background noise, divergence after double talk);
- switching characteristics of non-linear processor or centre clipper (NLP);
- variation of receive signal levels;
- echo loss and switching under double talk conditions with different signal level combinations.

If terminals are provided, the echo characteristics should be measured under single and double talk conditions using the same procedures and additional tests as described below. If handsets or headsets are used the environmental conditions should be modified:

- using a standard position relative to the HATS (normal telephone position); and
- in case of handset use, placing the handsets on a hard surface (transducers down) for echo measurements without double talk.

For hands-free use the tests should be carried out varying the acoustic environment.

10.2.6 Quality of Background Noise Transmission

Tests for evaluating the quality of background noise transmission will should be carried out by:

- simulating realistic background noise conditions;
- evaluating the performance of comfort noise adjustment (spectral, level).

10.2.7 Documentation of Test Results

The test results should be described and documented including the following:

- description of measurement setup;
- description of manufacturer settings;
- the test plan;
- results of instrumental assessment using ITU-T Recommendation P.862 [13] (electrical access)/TOSQA (acoustical access);
- results of correlation between instrumental and auditory assessment;
- processed speech samples (more than one language).

10.3 The VoIP Reference Point concept

The concept of a VoIP reference point is to specify an access point on the IP network whose characteristics are precisely controlled, like the ISDN on the switched network.

As shown on figure 19, the VoIP reference point is an interface between an IP domain and an analogue electrical domain. It is a device making the connection between IP packet flows and analogue speech signals.

This reference device is characterized by:

- compliance with specified codec implementations;
- specified amplitudes for speech signals;
- control of the processing delays.

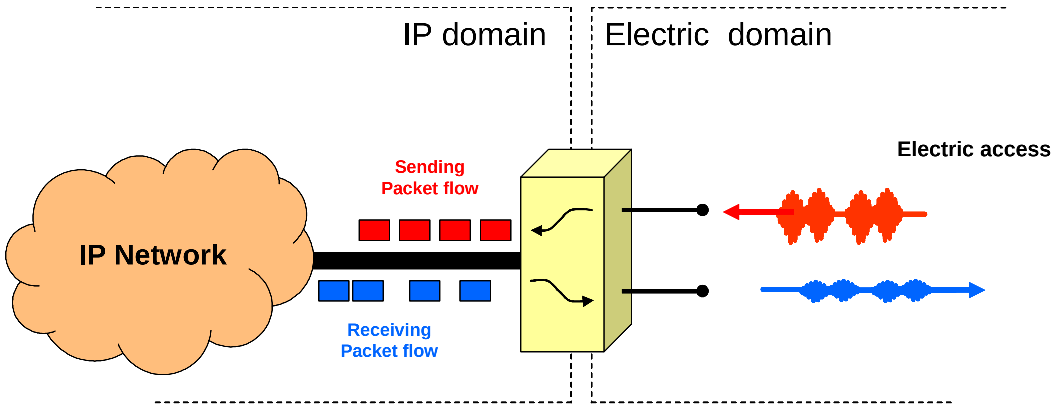


Figure 19: Concept of the VoIP reference point

The use of a reference point as a reference gateway makes possible the performance determinations of IP terminals.

Figure 20 shows the implementation of a VoIP reference point for the characterization of IP terminals. A call is establish between the reference device and the system to characterize allowing to an external measurement device to perform analysis.

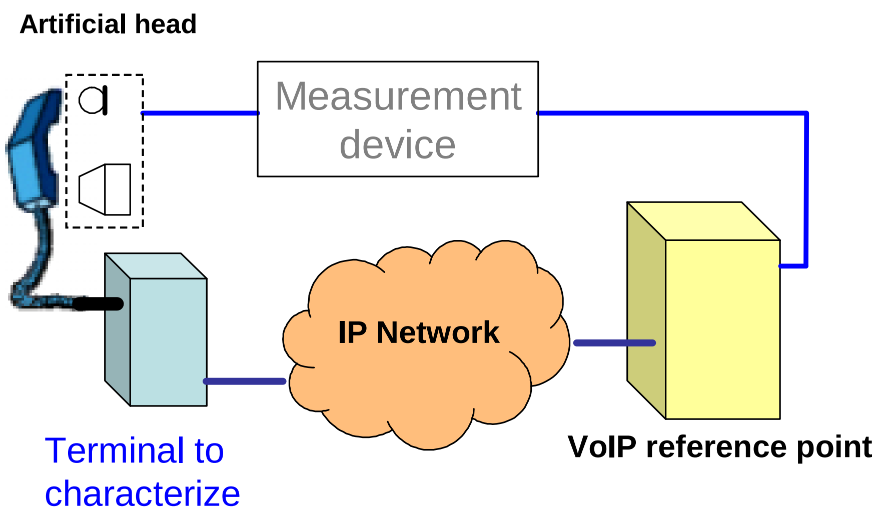


Figure 20: Test configuration using a VoIP reference point

11 Future work and unanswered questions

Work on quality of service and network performance for speech over packet networks is still in its infancy.

At the top level, the approach to end-end quality assessment developed for narrow band services based on MOS scales and R-factors needs to be revised as it does not necessarily carry across from narrowband to wideband.

In terms of network performance, much more research and accumulation of practical results is needed on the jitter performance of networks and its relationship to traffic congestion. For the NGN, limits to jitter need to be specified.

Then if performance is to be "guaranteed" or predicted, much more information is needed on how to design networks to achieve performance within specific levels of jitter.

In terms of services, if the NGN is to provide "packet pipes" then the dimensions of the pipes need to be standardized into a suitable range so that different interconnected networks can support compatible pipes.

In terms of the design of networks to achieve a high quality of service, it is essential that more information is given on the traffic types and volumes so that appropriate quality enhancement techniques can be chosen. It is particularly essential that the relationship of the NGNs to basic Internet traffic be clarified, as this will determine the efficacy of simple techniques such as prioritization.

If prioritization is to be used, then the NGNs need to standardize the DiffServ code points in terms of different classes of traffic and network performance.

Where networks are interconnected consideration needs to be given to planning for network performance and to the way in which impairments accumulate or UNI-UNI performance limits are apportioned across different networks. For example should there be an assumption that traffic congestion is random or correlated between different networks?

Firewalls and NATs and their traversal need further study to determine their effects on network performance parameters.

A great deal has been achieved in terms of terminal design and algorithms that tolerate jitter, but, whilst they may be effective in many cases, more results need to be accumulated before the effects can be quantified and used in calculations of end-end Quality of Service. In other words, the relationship between network performance and end-end quality is largely unknown in quantitative terms for modern codecs. This relationship needs much further work.

In terms of the customer, there is a need to research whether customers prefer a "guaranteed" minimum level of quality but with the possibility of access denial, or whether they would prefer reduced quality and lower probabilities of blocking. In terms of the approach, there is a need to consider whether blocking is a congestion or a quality of service measure.

Terminal design and PC configuration need further consideration to find ways to ensure that the best performance is achieved and that the set-up complexity is minimized for the human user.

Thus much further work is needed to establish an adequate framework for speech quality on IP-based networks.

History

Document history		
V1.1.1	June 2005	Publication