

Real-time communication over IP

Teletronikk

Volume 102 No. 1 – 2006
ISSN 0085-7130

Editor:

Per Hjalmar Lehne
(+47) 916 94 909
per-hjalmar.lehne@telenor.com

Editorial assistant:

Gunhild Luke
(+47) 415 14 125
gunhild.luke@telenor.com

Editorial office:

Telenor R&D
NO-1331 Fornebu
Norway
(+47) 810 77 000
teletronikk@telenor.com
www.teletronikk.com

Editorial board:

Berit Svendsen, Vice President Telenor Nordic
Ole P. Håkonsen, Professor NTNU
Oddvar Hesjedal, CTO Kyivstar GSM
Bjørn Løken, Director Telenor Nordic

Graphic design:

Design Consult AS (Odd Andersen), Oslo

Layout and illustrations:

Gunhild Luke and Åse Aardal,
Telenor R&D

Prepress and printing:

Rolf Ottesen Grafisk Produksjon, Oslo

Circulation:

3,600

Networks on networks

Connecting entities through networks – in technological, societal and personal terms – enables telecommunication. Networks occur on different levels, form parts of larger networks, and exist in numerous varieties. The artist Odd Andersen visualises the networks on networks by drawing interconnected lines with different widths. Curved connections disturb the order and show that networks are not regular but are adapted to the communication needs.

Per H. Lehne, Editor in Chief

Contents

Real-time communication over IP

- 1 Guest editorial – Real-time communication over IP;
Trond Ulseth and Finn Stafnsnes
- 3 Real-time communication over IP networks;
Trond Ulseth and Finn Stafnsnes
- 23 VoIP – Regulatory aspects from a Norwegian perspective;
Willy Jensen
- 27 Business models for broadband telephony;
Bjørn Are Davidsen and Finn Tore Johansen
- 40 Telephone Number Mapping (ENUM) – A short overview;
Trond Ulseth
- 43 Multimedia over IP networks; *Andrew Perkis, Peter Svensson, Odd Inge Hillestad, Stian Johansen, Jijun Zhang, Asbjørn Sæbø and Ola Jetlund*
- 54 Peer-to-peer IP telephony;
Paal Engelstad and Geir Egeland
- 65 Voice over IP in the context of 3G mobile communications;
Inge Grønabæk
- 82 Voice over WLAN (VoWLAN) – A wireless voice alternative;
Trond Ulseth and Paal Engelstad
- 97 MMoIP – Quality of service in multi-provider settings;
Terje Jensen
- 119 VoIP speech quality – Better than PSTN?
Trond Ulseth and Finn Stafnsnes
- 130 Security issues in VoIP;
Judith Rossebø and Paul Sijben

Special

- 149 The Telenor Nordic Research Prize 2005;
Terje Ormhaug
- 150 The Unpredictable Future: Personal Networks Paving Towards 4G;
Ramjee Prasad and Rasmus L. Olsen

161 Terms and acronyms in Real-time communication over IP

Guest Editorial – Real-time communication over IP

TROND ULSETH AND FINN STAFSNES



Trond Ulseth is Senior Research Scientist at Telenor R&D



Finn Stafsnæs is Research Scientist at Telenor R&D

When these words are written it is almost 130 years since Alexander Graham Bell patented a device converting speech into electrical signals that could be transmitted to a distant receiver; the telephone. On 10 March 1876 Mr. Bell spoke through the device to his assistant, Thomas A. Watson in the next room, the famous first words, “Mr. Watson – come here – I want to see you.”

An unknown businessman made another famous statement a few years later; “It is a scientific toy. It is an interesting instrument, of course, for professors of electricity and acoustics; but it can never be a practical necessity”.

Today we know that he was wrong, to the people in the industrialized world the telephone is a necessity, not a scientific toy.

Since the first demonstrations by Mr. Bell, there have been numerous new inventions and refinements to the system, both technically and commercially, but the main principles have remained unchanged; each connection is given a reserved amount of transport capacity; the circuit-switched network. Side systems have also been developed, such as support of supplementary services, IN services and emergency tele-communications systems where the call is set up to an emergency centre providing the centre with information about the address of the caller.

An all digital system, Integrated Services Digital Network (ISDN), introduced in the early 1990s has made it possible to offer new services such as videoconferencing in a public switched network. However, ISDN has not been a great success worldwide, and has not become the all-purpose network intended.

In the early 1960s a new concept, packet switching, was introduced. In packet switching the information is divided into packets that are sent independent of each other. Initially the primary packet switching applications were data communication, but the growth of the Internet has led to a number of new services offered on packet networks, e.g. interactive real-time communication – telephony and videoconferencing.

Although telephony may look simple at first glance, it is complex. In the beginning a lot of people considered voice and video communication as “just another data application”. They were wrong; real-time communication is different. Furthermore, developments

for more than one hundred years have resulted in a set of telephony services that users have been accustomed to and demand.

In spite of the ongoing hype, voice telephony over IP (VoIP) remains an early work in progress. There are also competing standards and a risk of different understanding of the standards which make interoperability demanding. Even though most vendors and operators today choose SIP as the protocol for VoIP provisioning, products from different vendors may not be compatible, so any company with a heterogeneous environment – that is, every large enterprise – may experience interoperability problems. Interop tests that aim at sorting out these problems are being held and will gradually sort them out, but it is unlikely that we have reached that point yet.

Users do not care about the technology, what matters are

- Functionality
- Quality
- Security
- Reliability
- Availability
- Cost.

However, the relative importance of these points varies widely between users and the actual use for the services. The main hurdles to VoIP might be speech quality and security.

Speech quality has traditionally been considered as a VoIP drawback. Many users have experienced reduced speech quality when the ADSL access line is used simultaneously for voice and Internet services. The quality may also be lower with increasing traffic in the core network, unless QoS mechanisms are implemented in the network. Although significant improvements have been made compared to the Voice over Internet products launched ten years ago, there are still hurdles. On the other hand, VoIP also gives an opportunity to provide better speech quality than traditional telephony by applying high quality speech codecs.

Among users there is a definite fear about VoIP security that has scared off many would-be customers. There are lots of threats to VoIP; from a bad guy hijacking the phone system, to eavesdroppers, to all kinds of nasty worms that use VoIP traffic to infect the entire network. Another negative effect to the

public is press reports about VoIP emergency calls that were routed to the wrong emergency call centres.

A term that was hype some years ago is 'Virtual telcos'. Essentially, it refers to telecommunications service providers who do not have their own network but instead choose to base their business model on using networks owned by others. It is a concept which is proving so attractive that new companies are appearing on a near daily basis, in both fixed and mobile markets.

The existence of virtual telcos has been made possible by a combination of regulatory interventions and an abundance of telecommunications capacity. This has been created through the combined effects of rapid technological advances and the ready availability of capital that turns these technical innovations into moneymaking infrastructure. The world's telecommunications capacity is growing even faster than its computing capacity. While computing capacity doubles every 18 months or so for the same price, according to Moore's Law, the information carrying capacity of optical fibre now doubles in just nine to 12 months, creating an overcapacity. This overcapacity enables the virtual telcos to offer services that are cheaper than those of traditional telcos.

Another issue is network cost. There are indications that network operators are reducing cost by replacing their circuit-switched networks with a packet-based networks.

Since the introduction of GSM in 1991 mobility has become important to users. In some countries the number of mobile telephony subscribers has become larger than the number of fixed network telephony subscribers. Mobile phones have made people accept lower speech quality than PSTN/ISDN could provide. On the other hand, it has also shown that people are willing to pay a higher price for services or functionality they consider important, e.g. mobility.

Many people who already have mobile phones have been willing to reduce the requirements to functionality, quality, security, reliability and availability if they can get fixed telephony at a lower cost, and switch to VoIP. VoIP may also add a mobility aspect to the fixed network telephony. You may bring your VoIP terminal with you, plug it into an Internet access anywhere in the world and keep your telephone number. However, this may be a concern to the regulators; at present this option limits the possibility to identify the caller's location in case of emergency calls.

Wireless LAN (WLAN/Wi-Fi) telephones may offer mobility – at home, at the company premises, or at access hotspots in hotels and other public areas.

A trend that has been discussed for a while is fixed/mobile integration. Standards organisations in the fixed and mobile worlds are now jointly developing multimedia standards allowing seamless services to be developed. The common term used is IMS (IP Multimedia Subsystem).

The number of traditional telephone network users is declining. On the other hand, the number of broadband customers is increasing. It is logical that people now are using their broadband access for more than surfing the web, e.g. to make telephone calls. We believe that new applications and services soon will become available making VoIP not only a low cost alternative to traditional telephony but the new way of communication.

No doubt, VoIP is the future. But all signs point to an evolution, not a revolution.

Trond Ulseth
Finn Stafnes

*Trond Ulseth is Senior Research Scientist at Telenor R&D. He obtained his MSc from the Norwegian University of Science and Technology in 1971 and has been working at Telenor R&D since 1972. The main topics of his work have been real-time communication (telephony and videotelephony). He has been active in standardisation work and chaired ETSI STC TE4 (Voice and Audiovisual Terminals) between 1990 and 1995. He has also participated in international research projects and is now working on projects on real-time communication over IP.
email: trond.ulseth@telenor.com*

*Finn Stafnes is Research Scientist at Telenor R&D. He received his MSc degree from the Norwegian University of Science and Technology (NTNU), department of Telecommunications in 1974 and joined Telenor (Televerket) R&D department in 1984. He has mainly been working with voice services and associated user- and network equipment during the evolution from the analogue PSTN via ISDN to VoIP and IP.
email: finn.stafnes@telenor.com*

Real-time communication on IP networks

TROND ULSETH AND FINN STAFSNES



Trond Ulseth is Senior Research Scientist at Telenor R&D



Finn Stafsnæs is Research Scientist at Telenor R&D

Standardisation work on VoIP protocols has now been going on for almost 10 years. The basic protocols can be considered as mature, but there is still a need for further developments. This article presents an overview of the two main groups of standards that specify VoIP protocols. Reference is also made to other articles in this issue of *Teletronikk*. New opportunities as well as unsolved issues are highlighted and future trends are indicated.

Introduction

Voice over IP (VoIP) has been characterised as the largest change to the telecommunication industry since the introduction of the Global System for Mobile Communication (GSM) 10 years ago. This statement is partly wrong. VoIP could be deployed as an alternative way to provide telephony with very little impact on the user experience. However, VoIP is an element of the transition from circuit-switched technology to packet-switched technology for all kinds of telecommunication applications including interactive real-time communication (i.e. telephony). The term Next Generation Network (NGN) is frequently used to describe this emerging network offering data, speech and real-time video on the same network.

Another important element contributing to VoIP deployment is the evolution of the DSL technology offering high-speed access to Internet.

Both proprietary and standardised VoIP products have been available for almost a decade. New small operators have seen this technology as a tool to enter a market that so far has been dominated by large operators. Now these large operators are offering VoIP products in competition with their own circuit-switched based telephony services. VoIP has also been an interesting technology for enterprise networks where the PABX is replaced by a VoIP system implemented on the company LAN.

VoIP will co-exist with telephony over circuit-switched networks for a long period. Solutions for interworking between telephony on circuit-switched networks and VoIP are therefore important.

Most VoIP networks are based on a client-server architecture. In March 2003 a company called Skype introduced peer-to-peer telephony. Skype has become popular among some groups of users, particularly because calls between Skype users connected to the Internet are free¹⁾.

Telephony services on circuit-switched networks have been standardised by ITU-T, supported by regional organisations such as ETSI. Another standardisation organisation, IETF, has carried out the standardisation of IP-based networks and applications. IETF has however limited their activities to technology and network mechanisms, while ITU-T (and the regional telecommunications standardisation organisations) also address user quality perception, reliability and operational issues. Both ITU-T and IETF are standardizing VoIP protocols. Proprietary solutions are also available.

From a technological, commercial and regulatory point of view VoIP is complex, for users and implementers alike. This article discusses the technological issues of VoIP and points to other relevant articles in this issue of *Teletronikk* presenting more details on some aspects.

What is VoIP (MMoIP)?

A simple answer to this question could be;

VoIP is voice communication over a packet-switched network using the Internet protocol (IP).

However, this is not the full answer. The way media is transported is only a part of the picture. Additional components are required in order to offer the users the desired functionality.

To many the terms 'the Internet' and 'IP' are two synonyms. This is partly true; IP is short for Internet Protocol. However, Voice over IP and Voice over the Internet are not synonyms. Voice over the Internet (Internet telephony) is any voice communication over the Internet; a public network that can be accessed worldwide. Voice over IP is voice communication over an IP network that can be a company LAN or the network of an ISP.

1) Except for the Internet access cost.

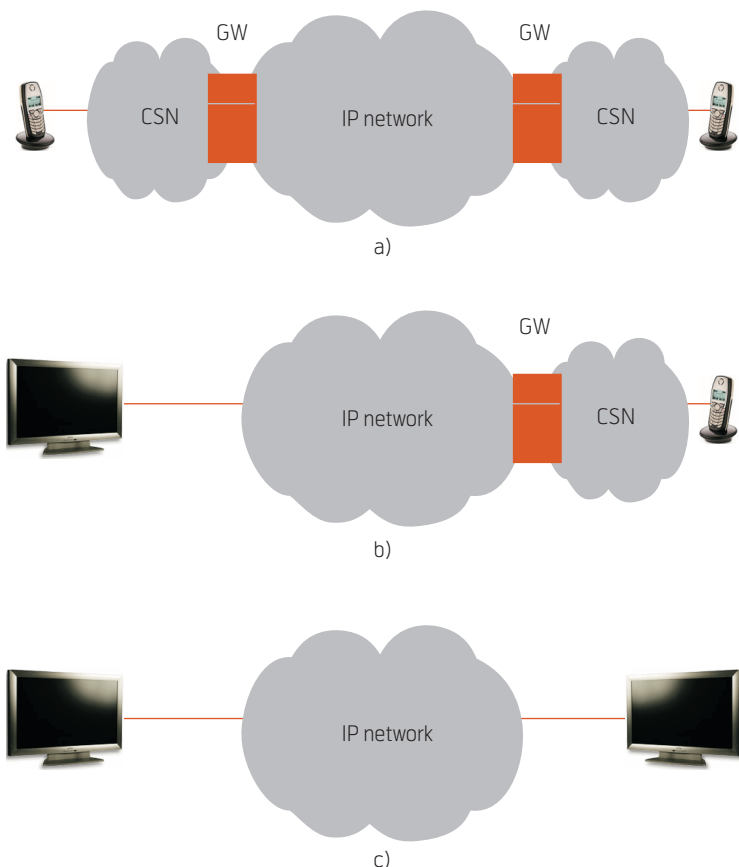


Figure 1 VoIP scenarios

Like many new technologies, VoIP was originally considered a cute novelty. But it is gaining momentum as a viable technology. There are several advantages, both in the home and in business, over traditional PSTN networks. A few include

- **Cost Reduction** – VoIP reduces the amount of network hardware needed (infrastructure overhead) by converging voice and data networks. Network efficiency improves with shared equipment. Excess bandwidth, rarely exploited, can be fully utilized.
- **Simplicity** – A single piece of equipment may support both voice and data communications. Less hardware means less cost. However, there are usability issues that may make dedicated pieces of equipment (e.g. a telephone) preferable.
- **Advanced Applications** – Like PSTN, basic telephony is the core element of VoIP. However, because VoIP uses a packetized digital format, the possibility for advanced multimedia (and multiservice) applications is limitless. A few possible applications include: Web-enabled call centers, collaborative white boarding, remote telecommuting, and personal productivity applications.

Although telephony may look simple at first glance; it is complex. Developments for almost one hundred years have resulted in a set of telephony services users have been accustomed to and demand. As VoIP systems have been developed and deployed, experience has revealed many features of circuit-switched networks which the initial versions of VoIP systems did not offer. In recent years a lot of work has been addressing functionality and features reproducing those of the circuit-switched networks telephony service.

This is important work ensuring that VoIP can provide a satisfactory replacement for circuit-switched telephone networks. However, VoIP can provide additional services to those of the circuit-switched network. The Internet environment is different from telephone networks in two fundamental ways. First of all, it is not limited to a single form of communication. Internet end systems can simultaneously be reading e-mail, browsing web pages, and sending instant messages, as well as communicating by voice. Secondly, the Internet is decentralized. Any end system can (in principle) communicate with any other; intermediate devices are only necessary if they provide some service to the end systems. As a consequence, VoIP can provide fundamentally new services, over and beyond those available in circuit-switched networks. An obvious example is adding video and/or data to the voice connection (Multimedia over IP – MMoIP). Simultaneous audio and video has been offered in ISDN, but communication over IP networks offers better flexibility as well as the possibility to use cheaper terminals.

Although the enthusiasts have declared that the next year is the ‘Year of VoIP’ since late 1990s, VoIP deployment so far has been slower than predicted. It is therefore clear that VoIP and telephony over circuit-switched networks (CSN) will co-exist for a long time. From a technical point of view VoIP will not be a single service or application. Figure 1 describes three possible VoIP scenarios;

- **Item a)** describes a scenario where VoIP technology is used in the transit connection between two CSN networks.
- **Item b)** describes a scenario where a connection between a VoIP user and a user connected to CSN is established.
- **Item c)** describes a connection between two VoIP users. The users may be connected to the network of the same VoIP service provider, or to separate VoIP service providers using different IP networks. In the latter case some interworking functionalities might be required.

Scenario c) is of course the ultimate scenario, and the main topic of this article. The IP part of scenario b) could be considered similar to scenario c); a Gateway could be considered as a terminal seen from the IP network point of view.

Wireless Local Area Network (WLAN) is a technology in great market growth, and new deployment scenarios and usage areas for WLAN appear at increasing frequency. Concurrently the technology is evolving in a rapid manner, constantly expanding its associated features. A natural extension to traditional data communication over WLAN is voice or multimedia communication over WLAN (VoWLAN). An overview of VoWLAN is presented in an article by Ulseth and Engelstad [1] in this issue of *Teletronikk*.

The history

The first experiments on voice communication over the Arpanet/Internet²⁾ were carried out late 1973. In the early 1990s an overlay network, Mbone, was used for conferencing applications. Mbone (Multicast Backbone) is a network of multicast enabled hosts that use the Internet as transport mechanism. In the Internet community Mbone has been used to broadcast events and to set up large real-time multi-party conferences. In March 1992 the 23rd IETF meeting was broadcast using Mbone. Other broadcast events are the NASA launches of space shuttles.

There have also been research activities and experiments on Mbone based multimedia conferencing. A project that belonged to the European Union ESPRIT programme was MICE (Multi-media Integrated Conferencing for Europe). This project ran from 1992 to 1995. Telenor R&D was one of the participants of the project.

Follow-up projects were MERCI (Multimedia European Research Integration) and MECCANO (Multimedia Education and Conferencing Collaboration over ATM Networks and Others). The MECCANO project concluded in 2000. Both ITU-T and IETF had at that time initiated standardisation work on standards for real-time communication over IP networks.

Early 1996 a company named Vocaltec launched the first VoIP service. The same year ITU-T published the first version of ITU-T Recommendation H.323 [2] specifying audiovisual communication over LAN that provided a non-guaranteed quality of service. Later versions have a more generic title; 'Packet-based multimedia communication systems'. At the

time being version 5 of the Recommendation is in force, a version that was published in July 2003.

IETF began its standardisation work on SIP (Session Initiation Protocol) in 1995. The first version of SIP was approved in March 1999. Version 2 of the protocol was approved in June 2002. It should be noted that while ITU-T keeps the old number when issuing a new version of a standard, IETF allocates a new number. Version 1 of SIP is therefore specified in RFC 2543, while version 2 of SIP is specified in RFC 3261 [3].

In the late 1990s several companies marketed VoIP solutions for corporate networks using proprietary protocols. Some of these companies are still marketing solutions based on these proprietary protocols along with standard based solutions. Most of the products on the market are however standards based. An important milestone is the 3GPP decision to use the SIP protocol in the IP version of third generation mobile network. 3GPP has developed a multimedia and telephony core network architecture called IMS (IP Multi-Media Subsystem). IMS is adopted by ETSI TC TISPAN as a basis for their work on standards for the Next Generation Network. This decision will simplify the Fixed network/Mobile Convergence that has been a 'hot topic' for several years already. An article by Grønbæk [4] in this issue of *Teletronikk* discusses IMS and the use of SIP in 3rd generation mobile.

In recent years the standards organisations working on the VoIP protocols have been focusing on standards specifying additional functionality.

The standards

There are two standards organisations working on standards for real-time communication on IP networks;

- International Telecommunication Union – Telecommunication Standardization Sector (ITU-T)
- Internet Engineering Task Force (IETF)

Both these organisations have developed a set of standards specifying signalling protocols for VoIP, known as H.323 (ITU-T) [2] and SIP (IETF) [3]. Both H.323 and SIP specify the use of the IETF Real-Time Transport Protocol (RTP) [5] for media transport.

2) *Arpanet is the forerunner of Internet.*

ITU-T Recommendation H.323

ITU-T Recommendation H.323 belongs to a set of recommendations specifying protocols and other characteristics for real-time two-way audiovisual communications (ITU-T H.320-series of recommendations). The first version of H.323 was issued late 1996, and could be seen as supplement to ITU-T Recommendation H.320 [6] that specified audiovisual communication on ISDN. ITU-T recommendation H.323 refers to two other ITU-T Recommendations.

- ITU-T Recommendation H.225.0 [7], which specifies call signalling protocols and media stream packetisation;
- ITU-T Recommendation H.245 [8], which specifies protocols for capability exchange and management including media stream management.

The H.323 protocol stack is illustrated in Figure 2.

The standards identified above specify the basic VoIP protocols and functionality. A separate series of ITU-T recommendations, the H.450 series, specify the procedures and the signalling protocols for support of supplementary services in H.323 networks. Part one (H.450.1) [9] is generic, while the other parts address specific services. These are:

- Call transfer (H.450.2)
- Call diversion (H.450.3)
- Call hold (H.450.4)
- Call park and call pick-up (H.450.5)
- Call waiting (H.450.6)

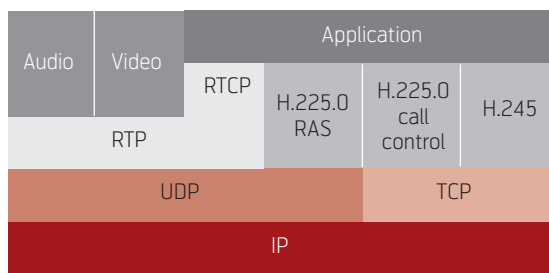


Figure 2 H.323 Protocol stack

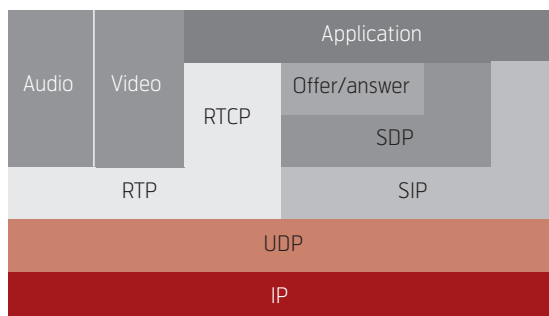


Figure 3 SIP Protocol stack

- Message waiting indication (H.450.7)
- Name identification (H.450.8)
- Call completion (H.450.9)
- Call offering (H.450.10)
- Call intrusion (H.450.11)
- Common Information Additional Network Feature (H.450.12).

Another set of Recommendations developed by ITU-T is the ITU-T Recommendation H.350 [10] series. These Recommendations specify a directory services architecture for multimedia communication and conferencing. Standardised directory services can support association of persons with endpoints, searchable white pages, and clickable dialling. Directory services can also assist in the configuration of endpoints and user authentication based on authoritative data sources, simplifying the management of large VoIP networks.

ITU-T has also developed a set of Recommendations specifying the functionality and protocol for data conferencing; the T.120 series of Recommendations [11]. These protocols can be used in a VoIP scenario enabling a MultiMedia over IP (MMoIP) scenario.

IETF Session Initiation Protocol (SIP)

Like ITU-T, IETF has also developed a set of standards specifying the protocols required to establish, modify, and terminate multimedia sessions. SIP is an ASCII-based, application-layer control protocol that has similarities with the Hypertext Transfer Protocol (HTTP) [12]. The Session Description Protocol (SDP) [13] is used to control media sessions. A mechanism by which two entities can make use of the Session Description Protocol (SDP) to arrive at a common view of a multimedia session between them is specified in RFC 3264 [14]. This mechanism is often referred to as the Offer/answer model. The SIP protocol stack is illustrated in Figure 3.

SIP specifies requests and responses. The SIP standard [3] defines six request methods:

- REGISTER is used to register a user agent (UA) at the location server
- INVITE is used to invite to a session
- ACK is used for acknowledgement
- CANCEL is used to cancel a request
- BYE is used to terminate a session
- OPTIONS is used to learn about the capabilities of a proxy or the other end-user agent.

Like ITU-T, IETF has also specified additional functionality to the basic SIP protocol. Both additional methods and header extensions are standardised. The additional methods standardised so far are

- INFO (RFC 2976) [15]
- PRACK (RFC 3262) [16]
- NOTIFY (RFC 3265) [17]
- SUBSCRIBE (RFC 3265) [17]
- UPDATE (RFC 3311) [18]
- MESSAGE (RFC 3428) [19]
- REFER (RFC 3515) [20].

At present more than 20 standards are approved specifying header extensions and events for SIP. There is also a number of other standards addressing SIP-related issues.

Another important issue is the decision of the 3rd Generation Partnership Project (3GPP) to use SIP as the voice signalling protocol for the IP version of the 3rd generation mobile network, discussed in [4].

VoIP architecture

Both H.323 and SIP are based on a client-server architecture as illustrated in Figure 4. The user agents (clients) register at the server of the service provider. It is however possible to set up a voice connection between two clients without making use of a server. To do so the IP address of the communication partner must be known.

H.323 architecture

In an H.323 network the server is called *Gatekeeper*. An H.323 gatekeeper controls the H.323 endpoints and its most important function is address translation between symbolic alias addresses and IP addresses. This way you use the caller name rather than the IP address currently allocated.

Important Gatekeeper functions are terminal registration and authentications. Endpoints attempt to register with a Gatekeeper at startup. This signalling is called RAS signalling (RAS – Registration, Admission, Status). The RAS protocol is defined in ITU-T Recommendation H.225.0 [7].

When a user wishes to communicate with another endpoint, request for admission to initiate a call using a symbolic alias for the destination endpoint, such as an E.164 address³⁾ or an e-mail address is sent to the Gatekeeper. If the Gatekeeper decides that the call can proceed, it returns a destination IP address to the originating endpoint. This IP address may not be the actual address of the destination endpoint, but it may be an intermediate address, such as the address of a proxy or another gatekeeper that routes call signalling further towards the destination.

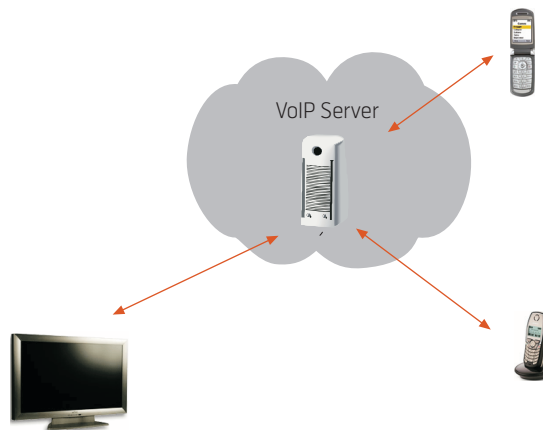


Figure 4 VoIP client-server architecture

The call control messages are then exchanged either directly between the endpoints or routed through the gatekeeper(s). The first case is called direct call signalling. The second case is called gatekeeper-routed call signalling. The Gatekeeper may also provide supplementary services such as call transfer, call forwarding etc.

The H.323 Recommendation introduces the term “zone”. An H.323 zone is a collection of all terminals, gateways, and MCUs managed by a single gatekeeper. A zone has only one gatekeeper. A call between endpoints belonging to two separate zones has to involve both Gatekeepers. The RAS protocol includes Gatekeeper-to-Gatekeeper signalling.

SIP architecture

The basic elements of a SIP system are user agents and *SIP proxy servers*. Like H.323 Gatekeepers, the SIP proxy servers route SIP requests towards their destinations. Proxies are typically collocated with a SIP registrar, which maintains a list of contact addresses for the users or accounts within the server’s IP domain.

The SIP user agent (UA) is software that is implemented in end-user devices to manage the SIP connection. User agents include endpoints such as IP phones, SIP media gateways, conferencing servers, and messaging systems.

The SIP Server functionality is divided into the following parts:

- SIP Registrar Server – handles registration messages and the location database.

³⁾ The term E.164 address refers to telephone numbers belonging to the international numbering plan for telephone systems defined in ITU-T Recommendation E.164.

- SIP Redirect Server – returns “contact this address” responses.
- SIP Proxy Server – forwards SIP requests and responses towards the destination.

Like in H.323, the user agents (endpoints) attempt to register with the registrar at startup. The SIP standard defines a Registrar Server as “a server that accepts REGISTER requests and places the information it receives in those requests into the location service for the domain it handles”. Authentications may also be carried out.

When making a call, the SIP proxies route requests to User Agent Servers⁴⁾ (UAS) and SIP responses to User Agent Clients (UAC). A request may traverse several proxies on its way to a UAS. Each will make routing decisions, modifying the request before forwarding it to the next element. Responses will route through the same set of proxies traversed by the request in the reverse order.

Peer-to-peer architecture

While the H.323 and SIP architectures are based on a client-server architecture, the peer-to-peer architecture has no centralized server. Peer-to-peer (P2P) technology was first deployed and popularized by file-sharing applications, enabling users to exchange audio, video and other types of data files. In 2003 a peer-to-peer application, Skype, was launched, and more companies are preparing the launch of similar services. The peer-to-peer architecture and other issues related to peer-to-peer telephony are discussed in an article by Egeland and Engelstad [21] in this issue of *Teletronikk*.

VoIP signalling

Transport of signalling

H.323 uses both unreliable (UDP) and reliable (TCP) signalling transport. The H.225.0 [7] RAS signalling (RAS – Registration, Admission and Status) uses unreliable transport (UDP). For the other signalling elements, i.e. Call set-up and the H.245 Control protocol [8], reliable transport (TCP) is used. The H.323 signalling is encoded using ASN.1 (Abstract Syntax Notation One). ASN.1 is a formal language for abstractly describing messages to be exchanged between applications.

ITU-T Recommendation H.235.0 [22] describes the H.323 security framework with common text and useful general information for the following parts in the H.235.x series of Recommendations. It applies to both H.323 signalling and media transport.

The SIP call control information messages can be sent using alternative transport protocols; UDP, TCP or SCTP⁵⁾. It is also possible to use TLS [23], in that case TLS is carried on top of TCP or SCTP.

Early implementations of SIP were often limited to use of UDP. Later implementations have frequently added the option to use TCP, which has benefits for the reliability of the SIP signalling. It may also help for NAT traversal. On the other hand, TCP is more complex and slower than UDP, so it has been less popular for low cost UA implementations. Even if a SIP element (server, client, UA or proxy) does not actively initiate SIP in TCP packets, it must be able to receive and respond to messages transported by TCP. This is a change from RFC 2543 (SIP version 1), but there are still many SIP devices around that are only capable of sending and responding to messages carried by UDP.

Whether the transport uses UDP or TCP, the main content in the SIP messages is normally sent in text format. Everything, except the authentication details, is readily readable for anybody who is able to get at the messages. When security of the signalling is considered important, TLS can be used. However, TLS is not yet available from most SIP server vendors, and probably even scarcer in UAs.

End-point registration

Both H.323 and SIP are based on a client-server architecture where user agents (clients) register at the server of the service provider.

In the Internet world e-mail addresses like ‘user@example.com’ and web URIs like ‘www.example.com’ are used. When an e-mail server sends messages to ‘user@example.com’ or a web browser sends requests to ‘www.example.com’, DNS servers are used to find the IP addresses where the messages or requests are to be sent. However, DNS can only be used to find addresses that do not change frequently.

VoIP users are often connected to access lines where the IP addresses change frequently on a more or less regular basis. The frequent change of IP addresses makes it impractical to use DNS servers to resolve IP-address queries for VoIP users.

⁴⁾ A User Agent (UA) consists of two parts, the UA Client (UAC), which initiates requests and the UA Server (UAS), which responds to requests.

⁵⁾ SCTP – Stream Control Transport Protocol, see RFC 2960.

One basic function of a VoIP server is therefore to keep track of the location where the VoIP users can be reached. The location information must generally be in a form that can be resolved using a DNS look-up or an IP identity.

A consequence of using a VoIP server to keep track of the location information of the users is that VoIP users in general can log into the service from any Internet access. This can be done by connecting the personal IP telephone (or other VoIP device) to an Internet access, or possibly even simpler by entering one's own VoIP address and password into an existing VoIP device that is already there. When the phone is connected or the user defined on the existing device, the service should be ready for use. VoIP can therefore be used for "nomadic mobility"; the service can be used from any location where there is Internet access. However, VoIP service providers may choose to restrict this possibility for regulatory or other reasons (e.g. caller location identification in an emergency situation).

In the present telephone network E.164 addresses consisting of digits only are used. A mechanism, ENUM, has been standardised that allows the translation of traditional telephone numbers into a format that can be used to store and retrieve Internet addressing information. An article by Ulseth in this issue of *Teletronikk* [24] gives an overview of ENUM. However, ENUM is not a mandatory mechanism, just an option.

In H.323 the address information is stored in the Gatekeeper. The endpoint sends a Registration Request (RRQ) to a Gatekeeper. The Gatekeeper responds with either a Registration Confirmation (RCF) or a Registration Reject (RRJ). The Gatekeeper address is either known to the endpoint through endpoint configuration, through an initialization file, or using the gatekeeper discovery mechanism defined in the Recommendation. One of the advantages of the Gatekeeper discovery mechanism is the ability to find an alternative Gatekeeper in fault conditions. These mechanisms are based on the RAS signalling. However, an endpoint should only register with a single Gatekeeper at a time.

It is possible to limit the duration of a H.323 registration by specifying this in the registration signals. Both the endpoint and the Gatekeeper may set this timer. Before the expiry of this timer a keep-alive RRQ may be sent. After the timer expiry a new, complete registration sequence has to be initiated.

Authentication as part of the registration process and other security related issues are addressed in ITU-T Recommendation H.325 [22]. Several scenarios are possible depending in the required security level.

In SIP the database where the address information is stored is called a location database. The main purpose of the SIP REGISTER request is to provide the necessary location details of where the SIP user can be contacted (Contact address) and for how long the Contact address shall be valid. The information is sent from the UA to the SIP registrar server, frequently co-located with a SIP proxy server. The duration of a registration can be suggested by the UA, but the Registrar can accept or modify the proposal. When the Registrar server has accepted the registration, the location details are stored in the Location database. If the user wishes to query what bindings that exist or the duration of the bindings, or the user wishes to remove (delete) registrations, specially formatted REGISTER requests can be used.

It is possible to have more registrations for one user. This possibility is used if a user would like to have an incoming call or other session presented at two or more destinations simultaneously. The task of diverting calls to more destinations is called Forking.

If the REGISTER request is accepted by the SIP server, the server returns an OK response to the UA. The response contains information of all current registrations for the user, not only the registration that was just done. In this way it is possible for the UA (SIP phone) to inform the user that if there is an incoming call, that call may also be presented at other destinations.

In common with HTTP, the possible response messages are given response numbers. The "OK" response is often referred as "200 OK". The first digit '2' in the response indicates that the request was successful.

If the request is not accepted straight away, there are many possible error messages that may provide information of what went wrong. In case of REGISTER request, the most common error message is "407 Proxy Authentication Required". This response contains a challenge and is used by the UA to authenticate the user; see below.

As stated earlier, the SIP signalling is very often sent in clear, readable text. It is easy for anybody who is able to gain access to a path for transmission to read the signalling, duplicate messages and construct malicious messages. To provide some protection against such misuse, most commercial SIP providers require users to authenticate themselves before the requests are accepted.

The normal authentication is based on HTTP authentication (RFC 2617 [25]), using a "Challenge – Response" principle. A challenge is sent from server

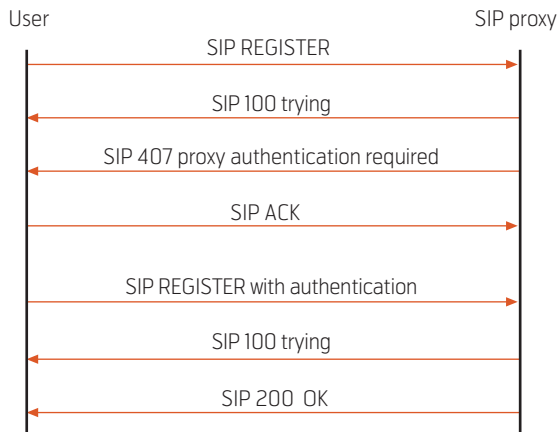


Figure 5 Normal sequence for registration, SIP

to UA. The Challenge, the user-ID and the password are used as input to an MD5 algorithm. The result from the processing is sent back to the SIP server. The SIP server uses the same input and the same algorithm to verify the response from the UA. In this way the password is not sent over the IP network. As the Challenge is different for every authentication sequence, it is not possible to use a replay of a previous authentication to get access to a SIP account. An article by Rossebø and Sijben in this issue of *Teletronikk* [26] discusses VoIP security aspects.

An example of a SIP registration sequence is shown in Figure 5.

Call set-up

The basic procedure for setting up calls in H.323 is to first send a RAS request (Admission Request – ARQ) to the Gatekeeper. The basic purpose of the request is to obtain the destination IP address. The Gatekeeper responds with an Admission Confirm (ACF) or an Admission Reject (ARJ) response. The Admission Confirm message contains the information required to call the desired destination endpoint or the Gatekeeper of the desired destination endpoint.

When an Admission Confirm response is received, the Call setup sequence is initiated. The H.323 call control messages are defined in ITU-T Recommendation H.225.0 [7]. The structure of the messages is similar to the structure of ISDN signalling, using reliable communication (TCP). The signalling messages may be routed via the Gatekeeper or sent directly to the destination endpoint. The information about the routing alternative to be used is included in the Admission Confirm message from the Gatekeeper.

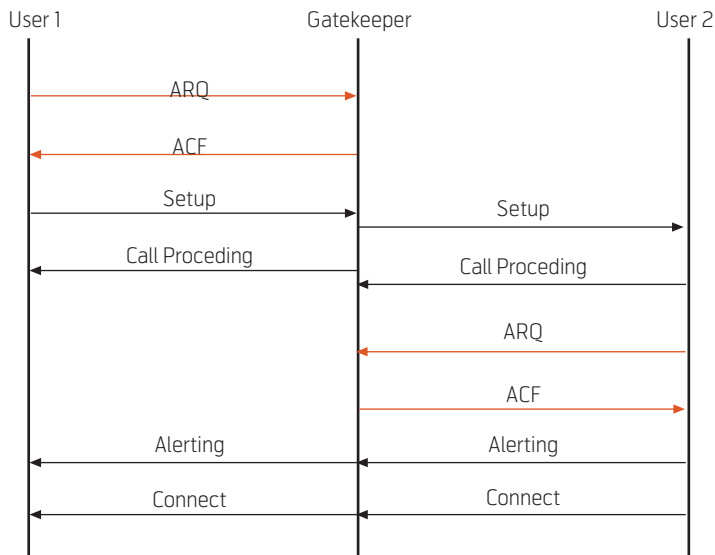


Figure 6 H.323 Gatekeeper routed call signalling, a single Gatekeeper

Signalling routed via the Gatekeeper results in more load on the Gatekeeper. However, Gatekeeper routed signalling can be used to support supplementary services and to produce call records for charging purposes.

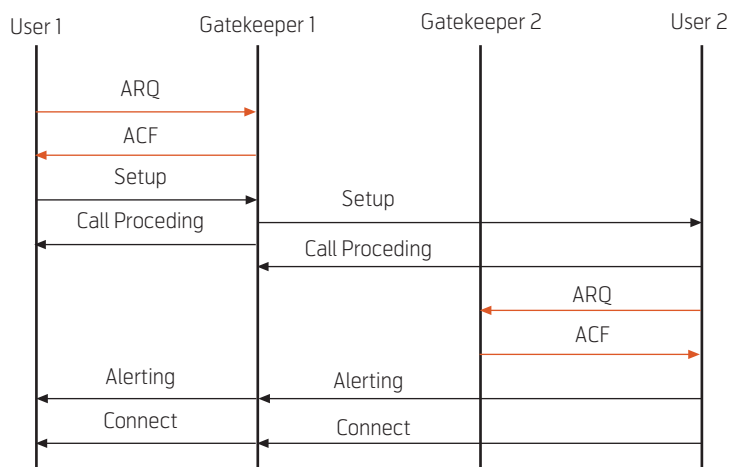


Figure 7 H.323 Gatekeeper routed call signalling, two Gatekeepers

Figure 6 shows a Gatekeeper routed call signalling sequence where both endpoints belong to the same zone (i.e. have same gatekeeper). When the endpoints belong to two different Gatekeepers, the call signalling may be routed via the Gatekeeper of the call initiating user as depicted in Figure 7 or via both Gatekeepers. The signalling procedures when the call is routed via both Gatekeepers are more complicated than the option depicted in Figure 7. ITU-T Recommendation H.225.0 [7] also defines RAS signalling between Gatekeepers.

To exchange capabilities and perform in-band negotiation at the start of or during communication, the protocols and procedures specified in ITU-T Recommen-

dition H.245 [8] is used. To do so a reliable H.245 control channel is established. Among the important functions are

- A master slave determination to avoid conflicts between the terminals involved;
- Capability exchange to agree on resources (e.g. audio and video codecs);
- Procedures for the opening and closing of logical channels which carry the audiovisual and data information;
- Mode (e.g. audio only, audio and video) control;
- Round-trip delay determination;
- Commands and indications for various purposes.

The basic method for setting up calls in SIP is to use the INVITE request. INVITE is sent from the calling party and is an invitation to the called party to participate in a session involving media that shall be exchanged between the participants. The media can be audio, video, other data or a combination of different media. In addition the media can be provided in many different forms, e.g. different codings for speech or video.

The INVITE request or subsequent signalling elements must contain the information that is needed to set up the session:

- The SIP address to the user(s) who are invited to participate;
- SIP address for the calling user;
- Contact address for the calling user, where and how the calling user can be reached;
- Different identifiers used to ensure that different requests are not mixed up;
- Detailed description of the session that the caller wants to establish (e.g. audio or video). The information is encoded according to a separate protocol, SDP (Session Description Protocol) [13], that is placed as “content” within the SIP message:
 - What kind of media to be used
 - Possible coding options for the different media
 - IP address and port number for each media stream.

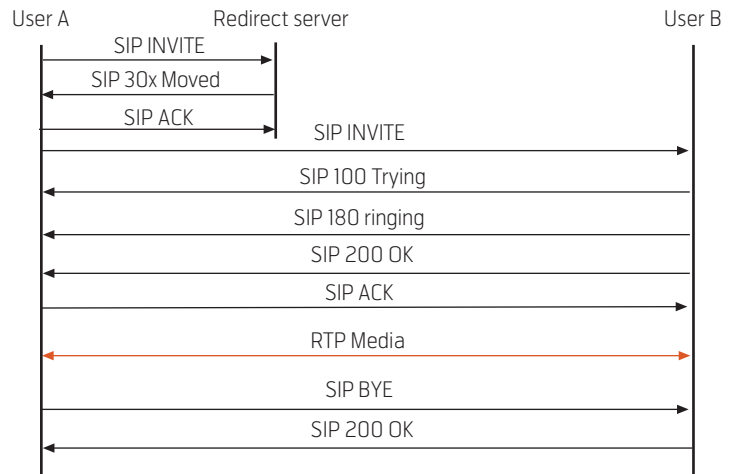


Figure 8 Signalling sequence, using Redirect server

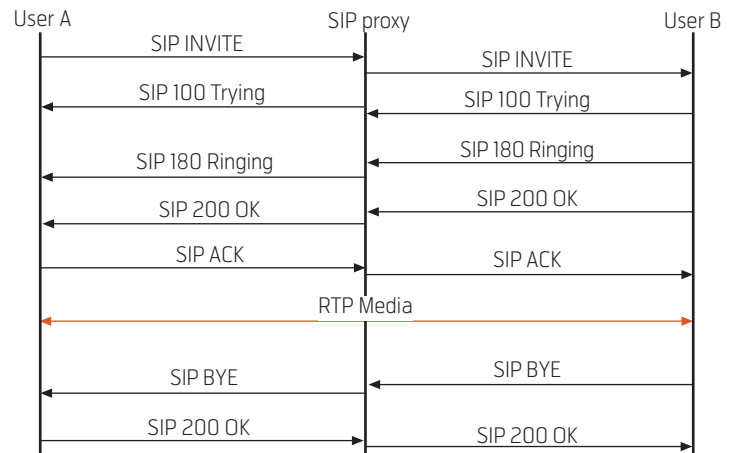


Figure 9 Signalling sequence, using Proxy server⁶⁾. Both users are known at the same Proxy server, all signalling is routed via Proxy

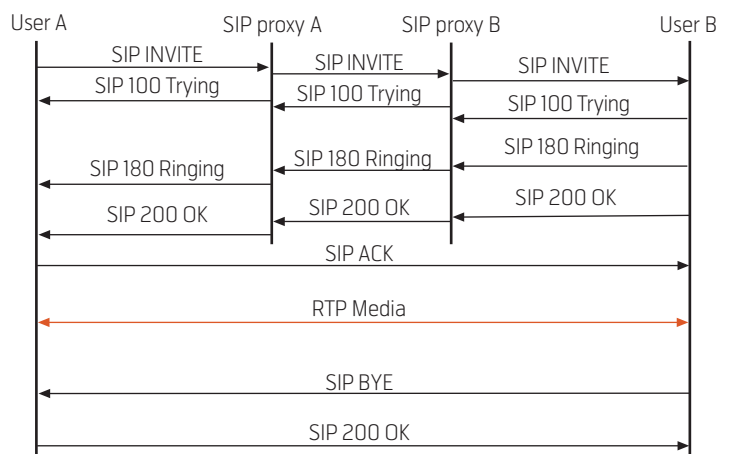


Figure 10 Signalling sequence, using Proxy server. The end users are known at different Proxies, only initial signalling is routed via Proxies

⁶⁾ The response “100 Trying” is normally generated as a provisional response from each signalling instance. Responses like “180”, “183” and “200” are normally generated at the destination and are routed backwards along the same path as the forward signalling.

There are basically two different types of SIP servers that can help in the process of call setup:

- Redirect servers
- Proxy servers

When receiving an INVITE request, a redirect server will return a redirection response (in the 300 series) with information where the called user can be contacted. Using this information, the UA can contact the called user without further involvement from the server.

The proxy servers, on the other hand, will try to forward the INVITE to the called user, without further involvement from the calling UA. If the proxy server does not have the called user in the local location database, it will forward the request to the SIP proxy server responsible for the SIP-domain that is found in the SIP address of the called user.

In both the above cases there will generally be one or more provisional responses (“100 Trying”, “180 Ringing”, “183 Call Proceeding”) and one response indicating success (“200 OK”). The calling UA shall confirm the connection with an ACK (acknowledge). If the call is not successful there will be one or more error messages coming back to the calling UA, in the 400 series, the 500 series or the 600 series depending on the reason for failure.

The calls are disconnected by sending a BYE request. The BYE can be sent from either party within the call. The other party accepts the BYE by returning “200 OK”.

In the proxy case, the INVITE request and the responses are routed through the proxy server. The ACK and signalling later during the session may go directly between the UAs, without going via the proxy. If a proxy wants to stay within the signalling loop it may specify the path for the signalling using a Record-Route header.

Figures 8, 9 and 10 show examples of such signalling. The figures also show differences in how ACK and BYE requests can be treated.

Description and negotiation for media exchange

When initiating a session the UA must specify a lot of details for the media and how they shall be used during the session. These details are provided using the Session Description Protocol (SDP) specified in

RFC 2327⁷⁾ [13]. SDP was made primarily to describe multicast sessions with one master controlling the set-up of the session and how media should be encoded and distributed. The protocol was not designed for negotiating media capabilities and options between peers.

The principles needed to be able to negotiate the session details between peers are covered in RFC 3264 [14]. This protocol, which is often referred to as ‘the Offer/Answer model’ is based upon SDP [13] and adds the negotiation mechanisms.

When setting up a call the caller must provide details for the session:

- What kind of session (e.g. speech or video and direction of media flow);
- How each media stream should be encoded, with acceptable options;
- Details for where the UA can be reached (IP address and port for each media stream).

In case of a SIP call, this information is sent as payload within a SIP message. Usually the caller includes the SDP content in the INVITE request message. When the callee receives the session description it must check the SDP content to see if it is compatible with the request or some of the options and return a response.

For many parameters, the response is an indication of whether the offer is accepted. For other parameters, e.g. speech or video encoding, the caller may provide a list of possible or acceptable encoding options. In such cases the called UA shall respond with its media capability options, taken from the list found in the invitation. In both cases the primary encoding scheme (preferred) is placed first in the list of options.

Usually the called UA includes the SDP response in one or more of the following SIP response messages: ‘180 Ringing’, ‘183 Session progress’ or ‘200 OK’.

With SIP there is no speech or video codec that all UAs must support. At least in theory, there is a possibility that a calling and called UA can find no common codec. In such cases the session invitation must be rejected, e.g. by sending the cause for incompatibility in a SIP error message, like the message ‘488 Not acceptable here’.

⁷⁾ A new version is under preparation.

In addition to normal call set-up, RFC 3264 also allows the UAs to ask the peer for media capabilities and also to change the session details during the session. Such change could involve adding or removing a media stream, putting a stream on hold, modifying the media encoding, changing the IP address or port where a stream should be sent to.

RFC 2327 and RFC 3264 specify the principles for session description and session negotiation. However, these RFCs do not specify the details for the different media encodings that can be used. These details are found in RFC 3551 [27] or other RTP profile documents.

Addressing principles

ITU-T Recommendation H.323 [2] has a flexible addressing system that allows both E.164 numbers and URLs.

Each H.323 entity shall have at least one network address. This address uniquely identifies the H.323 entity on the network. An H.323 endpoint may also have one or more alias address associated with it. An alias address may represent the endpoint or it may represent conferences that the endpoint is hosting. The alias addresses provide an alternative method of addressing the endpoint. These addresses include dialled digits (including private telephone numbers and public E.164 numbers), H.323 IDs (alphanumeric strings representing names, e-mail like addresses, etc.).

SIP uses URIs having the general format:
sip:user@host.domain, e.g. *sip:dduck@voip.online.no*
 The username can be a telephone number:
sip:85050000@voip.online.no

There are other possible addressing formats, most important the TEL URL
 General form *tel:<number>* e.g. *tel:85050000*

The TEL URL is specially devised to set up connections to the telephony network. Therefore options that are mainly applicable to addressing terminals within a telephone network can be included in the URL. These options cover global prefix, network prefix, ISDN sub address, digits to be dialled (as DTMF) after connection is established, just to mention a few.

This URL scheme has two other variants that can be mentioned:

- FAX URL
- MODEM URL

These variants are similar to the TEL URL. The URL starts with 'fax:' or 'modem', but are otherwise similar to TEL. They may use all the same optional information, but in addition they may add information that is applicable to fax connections or modem (data) connections.

Some VoIP terminals may only be able to enter destination addresses in the form of numbers (like telephone numbers). In this case two options can be used:

- Using the "phone number" as the user part of a SIP URI and add the host.domain part from the "own" service provider, e.g.
sip:21694922@voip.online.no. The calling user must leave it to the service provider to find out if the called user can be found in the providers own customer base or if the call must be forwarded to another provider or network.
- Using the TEL URL.

Those who are interested in addressing principles should also have a look at the overview presentation of ENUM by Ulseth in this issue of *Teletronikk* [24].

Media transport

Both H.323 and SIP use the same protocol for media transport, the RTP protocol [5]. Delay is a critical factor for speech communication. Retransmission of lost packets is therefore usually no alternative, and the UDP protocol is used instead of TCP. A typical media packet is illustrated in Figure 11.

Seven bits in the RTP header define the encoding of the payload. The payload type may be static; e.g. defined in an RFC, or defined dynamically through signalling. The static payload type of the speech coding algorithms that is normally used in VoIP, is defined in RFC 3551 [27]. RTP is also used for transport of video and end-to-end signalling (e.g. DTMF)⁸.

The payload size is a function of the coding algorithm used and the packet intervals. As an example, when a 64 kbit/s speech codec is used and the packet interval is 20 ms, the payload is 160 bytes. On the other hand, when an 8 kbit/s speech codec is used and the packet interval is 20 ms, the payload is 20 bytes. In the latter

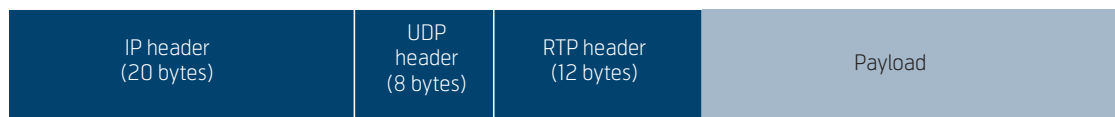


Figure 11 RTP packet example

scenario, the packet headers are twice the payload, in other words; there is a 200 % overhead.

Low bitrate codecs are often used because there is a link capacity problem in the connection. For the 8 kbit/s codec scenario described above, the required bitrate at IP level is 24 kbit/s; i.e. three times the encoded speech bitrate. The overhead might be larger due to the transmission technology used. As an example, an ADSL link using ATM technology to transport the IP packets adds more overhead. For the 8 kbit/s speech codec example the required ADSL bitrate when using 20 ms packet intervals is 63.6 kbit/s⁹⁾.

Increased packet interval may reduce the RTP packet overhead, but increase the end-to-end delay. Most applications use 20 ms packet intervals, as recommended in RFC 3551, but both TIA/EIA¹⁰⁾ and ETSI recommend 10 ms packet intervals. In the 8 kbit/s speech codec example the ADSL bitrate is 84.8 kbit/s when using 10 ms packet intervals.

One way to reduce the overhead problem is to use header compression. There are standards describing IP/UDP/RTP header compression, e.g. RFC 3095 [29] recommended by 3GPP. The compression algorithm may reduce the IP/UDP/RTP headers to 4 bytes or less. However, for the time being header compression is not frequently used in VoIP applications.

The effect of delay on perceived speech quality is discussed in an article [28] by Ulseth and Stafnes in this issue of *Teletronikk*.

RFC 3550 [5] also defines another protocol, the Real-Time Control Protocol (RTCP). RTCP is used to monitor the quality of service and to convey information about the participants in an on-going session. Each RTCP packet begins with a fixed part similar to that of RTP data packets, followed by structured elements that may be of variable length according to the packet type, but must end on a 32-bit boundary.

One important aspect of RTP is the dynamic allocation of port numbers. It is specified in the RTP protocol definition that RTP data should be carried on an even UDP port number and the corresponding RTCP packets are to be carried on the next higher (odd) port number. The only recommendation is that port numbers should be above 5000. This flexibility creates

problems for firewall and NAT transversal. These issues are discussed later in this article.

Speech quality

In the beginning VoIP had a reputation for being cheap and of low speech quality. It is still cheap, but there have been quality improvements. However, to achieve an acceptable speech quality a number of factors need to be controlled. An overview of these factors is presented in the article by Ulseth and Stafnes [28] in this issue of *Teletronikk*.

VoIP terminals

About 10 years ago, VoIP terminals were associated with a software application implemented on a PC. This is still an alternative, but there are others.

Many VoIP service providers offer a VoIP terminal adapter to which the user can connect his/her analogue telephone. The adapter may be a separate unit or integrated with other units such as an ADSL modem/router. One of the benefits of this solution is that the customers may continue using their existing telephone set.

The terminal adapter performs the speech encoding/decoding, the required signalling and packet network interfacing. The terminal adapter may also offer additional functionalities. One is QoS support. This may be achieved by routing data traffic via the adapter, giving priority to the speech communication over the data traffic.

Analogue telephone adapters (ATA) may be equipped with two or more analogue ports as well as ports supporting other functionalities. The analogue telephone interface needs to match the characteristics of the analogue telephone. These are different depending on the country of use. For Europe there is an ETSI Standard [30] specifying this interface. The standard is not developed to specify the characteristics of a VoIP terminal adapter, but the requirements to the analogue telephone interface are applicable to this adapter. Terminal adapters meeting these requirements interwork with analogue telephones meeting European type approval requirements.

There is an increasing number of stand-alone IP telephones on the market. They are usually equipped

8) When using low bitrate voice encoding, end-to-end DTMF signalling may be severely degraded if sent through the voice encoder/decoder. Therefore special formats for sending DTMF and other tones over RTP are defined.

9) Three ATM cells are required to transport the information. This includes also PPoE header, Ethernet header and encapsulation.

10) TIA/EIA are US industry organisations that develop telecommunications standards in the US.

with a display and feature keys that may be configurable. An IP telephone usually supports more functionalities than a telephone adapter. The speech quality characteristics of an analogue telephone and ATA combination are determined by the speech quality characteristics of the analogue telephone. These telephones are designed for narrowband¹¹⁾ speech communication. A designated IP telephone may be designed for wideband speech communication.

The third alternative, a PC or PDA with appropriate software, may use the built-in microphone/loud-speaker of the device, or separate handsets, headsets or microphone/loudspeaker combinations. The separate devices usually offer better performance than the built-in devices. The devices may be connected to the audio codec most modern PCs are equipped with or be equipped with a separate codec. In the latter case these devices are usually connected to the PC USB port.

The PC/PDA alternative introduces opportunities as well as limitations. The integration of speech communication with other functions is an obvious opportunity. The most obvious limitation is the bus structure and task priority mechanisms of the PC/PDA, which may introduce extra delay and jitter.

There are also wireless IP telephones communicating over WLANs. Some of these may also support mobile (GSM or UMTS) access. The article by Ulseth and Engelstad [1] on Voice over WLAN in this issue of *Teletronikk* describes handset terminals supporting this technology.

One of the advantages of VoIP is the possibility for the provider to upgrade the service by downloading new configuration or program files to the terminals.

Network QoS mechanisms

The Next Generation Network (NGN) is a packet-based network offering data, speech and real-time video on the same network. There are different requirements to the network performance; data applications require reliable communication where the packet transport time is not critical, while interactive real-time applications (e.g. telephony) are not so sensitive to packet loss but require short packet transport times.

To provide VoIP and MMoIP services in an NGN at an acceptable quality level network QoS mechanisms may be required. Most of the VoIP providers today

do not offer QoS functionality. The quality is still accepted by the users because there is an overprovisioning of transport capacity, and a lower quality is accepted when the product is cheaper. However, there are connection bottlenecks today (e.g. the access), and the traffic may increase more than the traffic capacity increases. In a longer term network QoS mechanisms are required.

An article by T. Jensen in this issue of *Teletronikk* [31] addresses challenges, service levels, mechanisms and potential gains related to proper managing QoS for multimedia-over-IP services.

Emergency telecommunications

Emergency telecommunications includes a broad spectrum of aspects related to the use of telecommunications services in emergency situations.

One aspect is to implement functions enabling identification of a caller's location. This is essential for e.g. the fire brigade, the police or an ambulance to find the actual site from where the emergency call has been made as quickly as possible.

Furthermore, nominated Emergency Control Centres of the emergency organizations deal with emergency calls from defined geographic areas, even when users from all districts call the same emergency number (e.g. 112). Emergency calls should be routed to and handled within, the appropriate Emergency Control Centre. There should be an unambiguous mapping between the location of the caller and the emergency control centre responsible for this area. For the time being VoIP services do not have the capabilities required to provide the customers with these functions. In the press there have been reports about emergency calls made by VoIP users that have been routed to emergency centres located in the area where the VoIP Gateway is located, not the area where the call is initiated.

A VoIP service may be nomadic; i.e. a VoIP service subscriber may log on to the service from any PC anywhere in the world as long as there is Internet access. This means that their location information may not be readily available.

The handling of these issues has been a hot topic among telecommunications regulators. An article by W. Jensen in this issue of *Teletronikk* [32] presents the approach chosen by the Norwegian regulator.

¹¹⁾ Descriptions of the terms narrowband and wideband speech communication are given in the speech quality article of this issue of *Teletronikk* [28].

In the United States FCC has placed an order where all operators that enable customers to receive calls from and terminate calls to the public switched telephone network (PSTN) are obliged to provide to the users the ability to make calls to emergency centres where location information is included. The information may be inserted manually by the users or included automatically. The order was effective by end of 2005. The location information may be entered manually into a database. This option is acceptable as long as the end user equipment is not moved from the location stored in the database.

One possibility to obtain location information in a nomadic use scenario is to use a DHCP¹²⁾ option specified in RFC 3825 [33]. The location information is obtained using principles described in RFC 3046 [34]. The transport of this information to a Gateway or an Emergency Control Centre is not specified in these protocols. Furthermore, mechanisms to route the emergency call to the correct Emergency Control Centre is not specified. Ongoing standardisation work in ETSI, IETF and ITU-T is attempting to solve these issues.

NAT and firewall transversal

The purpose of a Network Address Translator (NAT) is to translate IP addresses and port numbers from a private address space into public addresses when traffic flows from a private network to a public network. The main benefit of NAT is to extend the limited range of public IP addresses available.

NATs provide many benefits, but also have many drawbacks. The most troublesome of those drawbacks is the fact that they break many existing IP applications and make it difficult to deploy new ones. Guidelines have been developed [35] that describe how to build "NAT friendly" protocols, but many protocols simply cannot be constructed in accordance with those guidelines. Examples of such protocols include almost all peer-to-peer protocols, such as multimedia communications, file sharing and games.

A VoIP device that is connected at the private side of a NAT will know its private (or LAN) IP address and port number. When initiating a session the VoIP device must include the IP address and port number inside the call setup message that is hidden as payload inside an IP/TCP or IP/UDP packet. When the packet is forwarded out through the NAT, the IP address and port number are changed from private

to public values. However, in general the payload is not modified. Therefore the addressing information within the VoIP protocol elements will still contain private (LAN) values, which will be useless for subsequent routing of response messages or new calls back to the VoIP device.

The purpose of a firewall is to protect private network from being accessed by unauthorized sources. A firewall usually allows external (incoming) traffic only when it is initiated from the local network. Such behaviour may effectively block the possibility of incoming VoIP calls.

Today both residential networks and enterprise networks are protected by firewalls. NAT is also implemented in enterprise networks as well as in most residential networks.

There are several solutions to the NAT and firewall transversal problems. The article by Rossebø and Sijben in this issue of *Teletronikk* [26] presents an overview of the problems and solutions that may solve these problems. One of the solutions available, STUN specified in RFC 3489 [36] and NAT principles are discussed in the article by Egeland and Engelstad on peer-to-peer IP telephony [21] in this issue of *Teletronikk*.

An opportunity for new features and functionality

VoIP can be seen as a replacement for or an alternative to traditional circuit-switched telephony. Those who consider VoIP as a replacement for circuit-switched based telephony put a lot of effort on developing PSTN simulation or PSTN emulation services.

PSTN Simulation services provide PSTN/ISDN like services to advanced terminals (IP phones) or IP interfaces while PSTN emulation services provide PSTN/ISDN-like service capabilities using session control over IP interfaces and infrastructure.

The justification for this activity is to develop services and functionalities that are required by the regulators.

The standardisation of these services is carried out in the traditional telecommunications standardisation organisations such as ETSI and ITU-T, but documents describing the requirements of these services to the signalling solutions have been forwarded to IETF.

¹²⁾ DHCP (Dynamic Host Configuration Protocol) is the mechanism whereby the user is given an IP address when logging on an IP network.

IETF has also worked on protocols and procedures that provide functionalities similar to some circuit-switched network supplementary services.

IP communication adds more opportunities than just simulating or emulating PSTN services. It enables the users to add video and data (txt, graphics etc.) to a speech connection creating multimedia services. It is possible to establish multimedia services in ISDN, but IP communication offers more opportunities and flexibility. During a videoconference it might be possible to retrieve information from another source or add more video connections to an established speech connection.

Both groups of standards described earlier in this article, H.323 [2] and SIP [3], may be used in such communication scenarios. The starting point for the H.323 standard was audiovisual communication. The standard offers a lot of support to audiovisual communication, particularly through the H.245 [8] control part. The support of data is standardised in the ITU-T T.120 series of standards [11].

SIP does not have this functionality, but most of the video related functionalities are standardised in other IETF documents, particularly the RTP profile documents for the video coding algorithm used.

Interworking between Multimedia over IP and multimedia over ISDN is simpler when using H.323 than when using SIP because the H.323 protocol is based on the same principles as the ISDN-based H.320 protocol [6].

The IP technology enables the integration of real-time communication with other Internet-based applications. IP telephony users may call any number in an e-mail or Web application with one mouse click. The structure of the SIP protocol is similar to the http protocol [12], which simplifies the interworking between real-time communication and data communication.

So far, few attempts have been made at standardising or providing a standardisation framework for new features and functionality linked to VoIP. One of the few examples of such activity is Presence and Instant Messaging (IM).

In RFC 2779 [37] Presence and Instant Messaging are described as follows.

Presence is a means for finding, retrieving, and subscribing to changes in the presence information (e.g. "online" or "offline") of other users. Instant messaging is a means for sending small, simple messages that are delivered immediately to online users.

Some people may ask "Instant Messaging – isn't it the same as SMS in the mobile world?" The answer is "Yes, it is similar, but not identical". First, Instant Messaging messages can be sent to online users only, while SMS has a store and forward functionality. To learn whether the other user is online, the presence functionality can be used. Second, Instant Messaging may be used in multiparty communications.

There are stand-alone Instant Messaging services available on the market (e.g. AOL, Microsoft MSN and Yahoo). Instant Messaging may be linked to VoIP applications. Microsoft MSN offers a solution combining IM and VoIP. In IETF there is a separate working group, WG SIMPLE (SIP for Instant Messaging and Presence Leveraging Extensions), developing standard extensions for the transport of Instant Messages in SIP [38] and SIP support of presence [39].

Another SIP feature is third party call control. It is the ability of one entity to create a call in which communication is actually between other parties. Example application is transcoding, not only between two incompatible speech coding algorithms, but also between different media such as speech and text (text communication for deaf people).

Triple play is a term that is frequently used at present. The term describes a bundling of voice, video streaming and Internet access. There are not necessarily any links between these three applications, triple play can be seen as a marketing term assuming that an integrated solution is beneficial for the customers. However, triple play may contribute to faster deployment of VoIP.

IP multicast is a technology delivering source traffic to multiple receivers. The technology could be used to enable multiparty conferences, and multicast was used to set up large conferences over the Mbone network 10 to 15 years ago. Both ITU-T and IETF are developing standards on multiparty conferencing. The H.323 standard specifies the use of a Multipoint Control Unit (MCU) that connects the conference participants to each other. IETF has also chosen to base their standards on the same approach. In the IETF terminology this unit is called Focus. The multiparty conferences in an IP environment are therefore similar to PSTN/ISDN multiparty conferences.

While the ITU-T standard is stable, IETF is still working on their standards.

Real-time collaboration where voice and audiovisual conferencing are supported by tools such as Instant Messaging, Group Chat, live Web conferencing,

Screen sharing or Document sharing, adds more functionality to office users. IBM, Microsoft, Novell, Lotus, SAP, and Oracle are among the vendors currently building real-time collaboration capability into their products. Real-time collaboration is expected to improve the efficiency of business users.

The combination of eCommerce and VoIP may also encourage users to switch from traditional PSTN telephony to VoIP. An important element might be voice recognition enabling users to use their voice to make the desired choice.

Interworking with other networks

Among the VoIP scenarios described in Figure 1 two involve the interworking between an IP network and a circuit-switched network. It is expected that VoIP will co-exist with traditional telephony provided on circuit-switched networks for several decades. To set up a connection between a user connected to an IP network and a user connected to a circuit-switched network, an Interworking Function (IWF), often called a Gateway (GW), is required.

One Gateway approach is to connect it to the ISDN BRI (Basic Rate Interface) or ISDN PRI (Primary Rate Interface). This approach is often used to connect a corporate local area network (LAN) to the circuit-switched network. A single gateway both converts and transmits signalling information and media information.

Another approach is to separate the signalling and media gateway functions as illustrated in Figure 12.

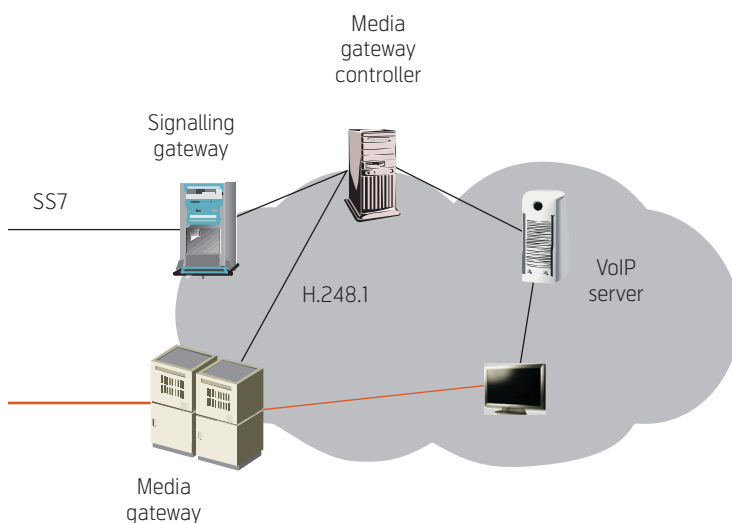


Figure 12 MEGACO architecture

This approach was proposed by ETSI EP TIPHON. Both IETF and ITU-T decided to develop a protocol based on this approach, and established a joint activity. However, due to different approval procedures the results, RFC 3525 MEGACO [40] and ITU-T Recommendation H.248.1 [41]¹³⁾, were not identical. Later ITU-T has published a new version of H.248.1 as well as a large number of packages that define additional functionality. IETF has not approved a new version corresponding to the latest version of the ITU-T Recommendation, nor documents corresponding to the additional functionality packages.

On the other hand, IETF has published an RFC that specifies an application of the MEGACO/H.248 Protocol for control of Internet telephones and similar appliances [42]. To achieve a high degree of interoperability and design efficiency in such end-user devices, a consistent architectural approach, a particular organization of Terminations and Packages, and a Protocol Profile have been described. The approach makes use of existing Protocol features and user interface related packages, and is thus a straightforward application of the MEGACO/H.248 Protocol.

The MEGACO/H.248 Protocol represents a single gateway control approach covering all gateway applications. This wide area of supported applications, along with its simplicity, efficiency, flexibility and cost-effectiveness, make MEGACO/H.248 a compelling standard for use in next generation networks. The standard is not tied to any particular call control approach and can be used both with H.323 and SIP. As shown in Figure 12, the architecture of MEGACO/H.248 includes the Media Gateway Controller (MGC) layer, the Signalling Gateway (SG), the Media Gateway (MG), and the MEGACO/H.248 Protocol itself.

Security and reliability

Users are today expecting telecommunications services that are secure and reliable. As with many other new technologies, VoIP introduces new security and reliability issues.

Security issues for VoIP include charging fraud, attacks on the infrastructure and protection of privacy. The article by Rossebø and Sijben [26] in this issue of *Teletronikk* presents an overview of threats to VoIP services and countermeasures to these.

Reliability aspects of VoIP services are addressed in an article by Johnson et al. [43]. Service reliability is

¹³⁾ The first version of the Recommendation was given the number H.248. Later H.248.1 is allocated to the basic Recommendation.

often described through downtime and defects per million metrics. A design criterion for the circuit-switched network is often referred to as the “five nines”, i.e. 99.999 % availability. This corresponds to about five minutes downtime per year.

An IP network designed for best-effort data communication will hardly meet these reliability requirements. The reliability requirements of this type of data traffic are lower in terms of short disruptions and delays caused by rerouting. The article states that router reliability must improve through better designs, higher hardware and software quality and fewer upgrades. Failure detection and recovery mechanisms are considered a necessity.

PCs have often been seen as the end-user device of VoIP. The reliability of a PC is often poorer than the reliability of designated terminals.

Finally, terminal adapters, IP telephones and PCs all require local power, while PSTN/ISDN terminals can be powered by the local exchange.

The reliability of VoIP depends on all these elements. It can be concluded that at present VoIP reliability is likely to be poorer than the reliability of telephony provided in a well-managed circuit-switched telephone network.

Future trends and conclusions

It is now (April 2006) more than ten years since the standardisation activities on VoIP solutions started, and almost ten years since the first standard (H.323 version 1) was published. Six years ago VoIP enthusiasts declared that next year would be the year of Voice over IP. At the same time there was also a lot of fuzz about ‘the new telcos’ that shortly would replace the traditional telcos. This has not (yet) happened. Why?

There are of course several reasons. The most important may be a lot of doubt about VoIP among the customers. Security, reliability and acceptable quality have been key issues. Reports in the press about emergency call problems contribute to this negative reputation.

A survey of IP telephony users conducted by Xelor Software [44] in the US finds widespread dissatisfaction with the technology’s QoS and reliability. More than 75 % of respondents who have deployed IP telephony have implemented some type of QoS procedure, the survey finds. However, more than 60 % of this group continue to receive complaints from users regarding voice call quality and reliability.

Residential market VoIP also depends on the deployment of broadband access. The asymmetry of the most popular broadband access technology today, ADSL, is a bottleneck for VoIP and MMoIP deployment.

The concentration gain in the links from edge router to DSLAM may also become a limitation when many DSL customers start using VoIP actively. Concentration factors between 20 and 50 are often used today. Access providers may have to adopt their products to the requirements of real-time communication.

Firewall and NAT transversal are challenges for the deployment of worldwide VoIP.

There are also commercial issues. The main market driver so far is cost, not features. The cost considerations are different for the business market and the residential market. For business users VoIP means cheaper infrastructure. Business users may therefore replace their enterprise telephone network with VoIP implemented on their local area network, and a gateway between the LAN and the public network. The infrastructure cost for residential users is however low, they have no infrastructure related cost benefit.

It is claimed that VoIP means the end of traditional pricing. The VoIP products available today are cheaper than PSTN/ISDN telephony. There are however indications that the PSTN/ISDN operators are modifying their pricing structure to combat the VoIP challenges. On the other hand, these operators are launching VoIP services in competition with their own telephony service.

There is a large number of VoIP providers in most western countries. In Norway the NPT has registered more than 20 providers. For the users, this large number of providers creates a complicated market scenario. Providers may offer non-standardised solutions. One example is Skype. The consequence is that the user can only communicate with Skype users or make use of additional services such as Skype in/Skype out, which are not free.

Another problem is the lack of interconnect agreements between VoIP providers. Again, a call between two users that are subscribing to two different VoIP providers may be set up via a circuit-switched network.

We cannot expect that the customers are able to understand all implications of incompatibility and interconnection problems.

Usability may also be a bottleneck for VoIP deployment. The set-up of a VoIP call is slower than the set-up of a PSTN/ISDN call. The users may be negative to this.

The PC is not always the optimal device for voice communication, which may influence the user experienced performance and usability.

Another question is whether future VoIP products will be based on the client-server architecture or a peer-to-peer architecture such as Skype. At present the most successful peer-to-peer IP telephony solution uses proprietary protocols. However, as pointed out by Egeland and Engelstad in this issue of *Teletronikk* [21], SIP might be used as the signalling protocol in a peer-to-peer IP telephony setting.

The VoIP potential so far is not realized in terms of new features. Relevant questions are:

- Will Multimedia over IP be a success?
- Will wideband speech replace narrowband speech?
- Will Instant Messaging and Presence be a success (in Europe)?
- Will mobility solutions and VoWLAN contribute to the deployment of VoIP?
- Will integration of voice and data (web applications) be a killer application?

Videotelephony and videoconferencing have not been a success so far. The use of MMS in the mobile market is increasing and may be an intermediate step towards mobile videotelephony. This trend may also help Multimedia over IP (videotelephony over IP) deployment. The customers may be used to live video over a telecommunication link, while there might be a preference for the larger screen of a PC or a videophone. It is also essential that the access link capacity is adapted to the requirements of two-way video communication.

Standardised wideband speech coding algorithms have been available for more than 15 years. Users prefer wideband speech for videotelephony and videoconferencing, but the use of wideband coding in telephony applications have been almost non-existent. Press reviews of Skype where wideband speech codecs are implemented, have however highlighted the superior speech quality. This is an indication that users may prefer the improved quality of the wideband codecs. Wideband speech puts some new requirements on the electroacoustic transducers,

and may therefore be an alternative when headset- or loudspeaker-based terminals are used.

The differences between Instant Messaging (IM) and SMS are described earlier in this article. People who are used to SMS may find IM less attractive. A survey has indicated that more than 40 % of the Internet users in the US also use IM [45]. The trend might be different in regions where the mobile (GSM) density is high compared to regions where mobile density is low. The inclusion of presence might make IM more attractive. It is therefore difficult to predict the long term popularity of IM/Presence in Europe.

VoWLAN speech quality has been reported as a problem area. The handover and roaming mechanisms are missing. VoWLAN cannot be considered as an alternative to mobile systems such as GSM or 3G, but rather a supplement to fixed network VoIP terminals like DECT in the PSTN/ISDN world. The VoWLAN equipment cost is still high. VoWLAN can therefore be considered as an alternative for selected areas of use, but will hardly contribute significantly to the deployment of VoIP for the time being.

As already indicated in this article, integration of voice and web applications is simpler in an IP environment, particularly when using the SIP call set-up protocol. It is likely that this combination will be a future success, but it cannot be considered as a VoIP killer application for the time being. There are however indications that something is happening. The combination of VoIP and IM/presence has already been addressed in this article. Another combination that has been addressed recently is integrating audio conferencing with Web Conferencing, giving participants in Web conferences a visual indication of who is speaking, as well as being able to dial out to new participants, mute lines and control volume.

The network operators may replace the existing circuit-switched network with a packet-switched network some time in the future. When this happens, the telephony services have to be IP-based, i.e. VoIP/MMoIP. However, VoIP is already offered, but for the time being as a low price alternative to traditional telephony.

So far the VoIP products offered are similar to circuit-switched based telephony, but not all of the functionalities are available. Some of the supplementary services in circuit-switched networks are not frequently used, but a few are a 'must' for many users. Not all of these are available for the time being, and some VoIP providers also have problems meeting some of the regulatory requirements.

People are used to be able to call any telephone subscriber worldwide from their telephone. This is not possible in VoIP today. Three groups of terminals are not interworking;

- H.323 terminals
- SIP terminals
- Terminals where proprietary protocols are implemented.

In enterprise networks, a number of H.323-based and proprietary solutions are implemented. One of the reasons why enterprises choose these solutions may be better support of video communication than the present SIP based products on the market.

Most of the providers of VoIP services to residential customers base their products on the SIP protocol. A lot of effort is made to secure that implementations from different suppliers interwork. Regulatory requirements may on a short term be an obstacle to VoIP growth, but may be beneficial on a longer term, improving the thrust to VoIP.

Fixed Mobile Convergence has been an issue for quite a while. The progress of VoIP and IMS in fixed networks together with the introduction of the IP technology (and IMS) in the mobile world as highlighted in the article by Grønbaek [4] would simplify this process. Personalisation making the user interface and available services independent of network accessed and terminals used would also contribute.

There are variations of the regulatory requirements to VoIP between different countries. Some countries have banned the use of VoIP services, other countries consider VoIP an Internet application that is not regulated at all. There is however a trend that VoIP will be regulated, and that the regulatory requirements are harmonised. This trend may affect the VoIP deployment. On a short term this may be a disadvantage, but the authors of this article believe that it is an advantage to VoIP deployment on a longer term.

It can be concluded that there are still hurdles, there is a need for more standardisation work, improved quality and reliability, but the future looks prosperous.

References

- 1 Ulseth, T, Engelstad, P. Voice over WLAN (VoWLAN) – A wireless voice alternative? *Teletronikk*, 102 (1), 82–96, 2006. (This issue)
- 2 ITU-T. *Packet-based multimedia communications systems*. Geneva, 2003 (ITU-T Recommendation H.323)
- 3 IETF. *SIP: Session Initiation Protocol*. 2002 (RFC 3261)
- 4 Grønbaek, I. Voice over IP in the context of 3G mobile communications. *Teletronikk*, 102 (1), 65–81, 2006. (This issue)
- 5 IETF. *RTP: A Transport Protocol for Real-Time Applications*. 2003 (RFC 3550)
- 6 ITU-T. *Narrow-band visual telephone systems and terminal equipment*. Geneva, 2004 (ITU-T Recommendation H.320)
- 7 ITU-T. *Call signalling protocols and media stream packetization for packet-based multimedia communication systems*. Geneva, 2003 (ITU-T Recommendation H.225.0)
- 8 ITU-T. *Control protocol for multimedia communication*. Geneva, 2005 (ITU-T Recommendation H.245)
- 9 ITU-T. *Generic functional protocol for the support of supplementary services in H.323*. Geneva, 1998. (ITU-T Recommendation H.450.1).
- 10 ITU-T. *Directory services architecture for multimedia conferencing*. Geneva, 2003 (ITU-T Recommendation H.350)
- 11 ITU-T. *Data protocols for multimedia conferencing*. Geneva, 1996 (ITU-T Recommendation T.120)
- 12 IETF. *Hypertext Transfer Protocol – HTTP/1.1*. 1999 (RFC 2616)
- 13 IETF. *SDP: Session Description Protocol*. 1999 (RFC 2327)
- 14 IETF. *An Offer/answer Model with the Session Description Protocol (SDP)*. 2002 (RFC 3264)
- 15 IETF. *The SIP INFO Method*. 2000 (RFC 2976)
- 16 IETF. *Reliability of Provisional Responses in the Session Initiation Protocol (SIP)*. 2002 (RFC 3262)
- 17 IETF. *Session Initiation Protocol (SIP) – Specific Event Notification*. 2002 (RFC 3265)
- 18 IETF. *The Session Initiation Protocol (SIP) UPDATE Method*. 2002 (RFC 3311)
- 19 IETF. *Session Initiation Protocol (SIP.) Extension for Instant messaging*. 2002 (RFC 3428)

- 20 IETF. *The Session Initiation Protocol (SIP) Reference Method*. 2003 (RFC 3515)
- 21 Egeland, G, Engelstad, P. Peer-to-Peer IP Telephony. *Teletronikk*, 102 (1), 54–64, 2006. (This issue)
- 22 ITU-T. *H.323 security: Framework for security in H-series (H.323 and other H.245-based) multimedia terminals*. Geneva, 2005 (ITU-T Recommendation H.235.0)¹⁴⁾
- 23 IETF. *The TLS Protocol Version 1*. 1999 (RFC 2246)
- 24 Ulseth, T. Telephone Number Mapping (ENUM) – A short overview. *Teletronikk*, 102 (1), 40–42, 2006. (This issue)
- 25 IETF. *HTTP Authentication: Basic and Digest Access Authentication*. 1999 (RFC 2627)
- 26 Rossebø, J Y E, Sijben, P. Security issues in VoIP. *Teletronikk*, 102 (1), 130–145, 2006. (This issue)
- 27 IETF. *RTP Profile for Audio and Video Conferences with Minimal Control*. 2003 (RFC 3551)
- 28 Ulseth, T, Stafnes, F. VoIP speech quality – Better than PSTN? *Teletronikk*, 102 (1), 119–129, 2006. (This issue)
- 29 IETF. *RObust Header Compression (ROHC): Framework and four profiles: RTP, UDP, ESP, and uncompressed*. 2001 (RFC 3095)
- 30 ETSI. *Access and Terminals (AT); Public Switched Telephone Network (PSTN); Harmonized specification of physical and electrical characteristics at a 2-wire analogue presented Network Termination Point (NTP)*. Sophia Antipolis, 2002 (ETSI TS 201 970 v1.1.1)
- 31 Jensen, T. MMoIP – Quality of service in multi-provider settings. *Teletronikk*, 102 (1), 97–118, 2006. (This issue)
- 32 Jensen, W. VoIP – regulatory aspects from a Norwegian perspective. *Teletronikk*, 102 (1), 23–26, 2006. (This issue)
- 33 IETF. *Dynamic Host Configuration Protocol Option for Coordinate-based Location Configuration Information*. 2004 (RFC 3825)
- 34 IETF. *DHCP Relay Agent Information Option*. 2001 (RFC 3046)
- 35 IETF. *Network Address Translator (NAT)-Friendly Application Design Guidelines*. 2002 (RFC 3235)
- 36 IETF. *STUN – Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)*. 2003 (RFC 3489)
- 37 IETF. *Instant Messaging / Presence Protocol Requirements*. 2000 (RFC 2779)
- 38 IETF. *Session Initiation Protocol (SIP) Extension for Instant Messaging*. 2002 (RFC 3428)
- 39 IETF. *A Presence Event Package for the Session Initiation Protocol (SIP)*. 2004 (RFC 3856)
- 40 IETF. *Media Gateway Control Protocol (MEGACO) version 1*. 2003 (RFC 3525)
- 41 ITU-T. *Gateway Control Protocol version 2.0*. Geneva, 2002 (ITU-T Recommendation H.248.1)
- 42 IETF. *Megaco IP Phone Media Gateway Application Profile*. 2001 (RFC 3054)
- 43 Johnson, C R, Kogan, Y, Levy, Y, Saheban, F, Tarapore, P. VoIP Reliability: A Service Provider’s Perspective. *IEEE Communications Magazine*, 42 (7), 2004.
- 44 [online] October 2005. URL: http://www.xelorsoftware.com/home/releases/10_27_05.html
- 45 [online] January 2006. URL: http://www.voip-weekly.com/features.php?feature_id=80

For a presentation of the authors, please turn to page 2.

¹⁴⁾ The ITU-T Recommendation H.235 (2003) content was reorganized into H.235.0 to .7 when revised in 2005.

VoIP – Regulatory aspects from a Norwegian perspective

WILLY JENSEN



Willy Jensen is General Director of the Norwegian Post and Telecommunications Authority

Although the regulation of Voice over IP (VoIP) services still raises some challenging issues in Norway, legal certainty is achieved to a great extent. This article outlines the regulatory initiatives taken by the Norwegian Post and Telecommunications Authority (NPT) regarding VoIP services. Three main categories of VoIP services are identified that may be treated differently for regulatory purposes. Focusing on the third category it is concluded that such VoIP services are subject to the same legal requirements as providers of traditional telephone services, but that temporary exemptions may be given from some of the obligations for an interim period. NPT has also concluded that such VoIP services are substitutable with traditional fixed telephony (PSTN/ISDN), and these services are therefore included in the Norwegian retail and wholesale markets for fixed telephony (Markets 1-6 and 8-10).

1 Introduction

VoIP services using ordinary telephone numbers with possibilities for any-to-any communication were introduced in the Norwegian market in the beginning of 2004. These services have been well received in the market and by the end of 2004 almost 50,000 end users subscribed to such services. The number of users is growing fast and there were about 110,000 subscribers as per 1 July 2005. Norwegian Post and Telecommunications Authority (NPT) estimates that there were at least 175,000 subscribers by the end of 2005 (about 8-9 % of the total number of fixed telephony subscribers in Norway). Other types of VoIP services are also used in Norway, e.g. the plain version of Skype. NPT has no information about the number of users of such services.

This article will focus on the regulation of VoIP services in Norway. Being a member to the EEA treaty, Norway has adopted the European regulatory framework for electronic communications into its national legislation. Norway has no specific *sui generis* legislation applicable to VoIP. Instead the general rules relating to electronic communications services apply. The most important rules regulating VoIP in this context are contained in the Electronic Communications Act, the Electronic Communications Regulations and the Numbering Regulations.

NPT is set to administer the rules of the said act and regulations, and to ensure that they are adhered to by the market itself. Following a public consultation¹⁾ NPT published a policy paper 15 April 2005 setting

out how certain VoIP services are regulated under Norwegian Law (the Policy Paper).²⁾

The Policy Paper is in line with the principles set out in the European Regulators Group's (ERG) "Common Statement for VoIP regulatory approaches" which was published 11 February 2005.³⁾ It has been important for NPT to adopt a regulatory approach that is consistent with the objectives of the European regulatory framework and will enable the greatest possible level of innovation and competitive entry in the market, whilst ensuring that end user interests are adequately protected.

It is worth mentioning that market regulatory issues regarding significant market power (SMP) are not dealt with in the Policy Paper. However, NPT has also considered whether VoIP should be included in the relevant markets for fixed telephony defined by the European Commission (Markets 1-6 and 8-10 in the Commission Recommendation on Relevant product and service markets).⁴⁾ The outcome of these considerations is briefly described in section 4 below.

2 General contents of the VoIP Policy Paper

Generally, the Policy Paper does not contain any *de lege ferenda* discussions on how VoIP services should be regulated in an ideal world, but merely points out how the current regulatory regime will be applied to such services. The Norwegian Ministry of Transport and Communications is currently analysing

1) The consultation is published (in Norwegian only) on NPT's website (<http://www.npt.no>).

2) The Policy Document is published (in Norwegian only) on NPT's website. An English text synopsis is also available.

3) The ERG document can be downloaded from the following web address: http://www.erg.eu.int/doc/publications/erg0512_voip_common_statement.pdf.

4) Commission Recommendation of 11 February 2003 (2003/311/EC)

whether the rules relating to, *inter alia*, VoIP services should be amended. However, it is too early to predict the outcome of these analyses.

The Policy Paper identifies the following three main categories of VoIP offerings:

Category 1 VoIP offerings which are not any-to-any communication enabled. Within this category, no gateway to the PSTN/ISDN or mobile networks exists, and hence no possibility to call or receive calls from traditional telephone services. An example of category 1 VoIP offerings is the plain version of Skype.

Category 2 VoIP offerings which are partly any-to-any communication enabled. Within this category, a gateway to the PSTN/ISDN or mobile networks exists, giving the possibility to *either* call *or* receive calls from POTS, but not to *both* call *and* receive calls to/from such services. An example of category 2 VoIP offerings is Skype Out.

Category 3 VoIP offerings which are any-to-any communication enabled. Within this category, a gateway to the PSTN/ISDN or mobile networks exists, giving the possibility to both call and receive calls from POTS. Most of the VoIP providers currently operating in Norway fall within this category.

NPT has concluded that category 3 VoIP services fall within the scope of the Electronic Communications Act, and that they are, if available to the public, publicly available telephony services (PATS).

NPT has yet to conclude whether VoIP offerings that fall under category 1 or 2 are within the scope of the Electronic Communications Act.⁵⁾ A national hearing on this subject was conducted in the autumn of 2005. Following this consultation, NPT will publish its view on the matter in the spring of 2006.

3 Regulation of category 3 VoIP services

Since VoIP services falling within category 3 are deemed as PATS services according to the NPT Policy Paper, the whole set of obligations relating to Electronic Communications Services (ECS) and

PATS under the Electronic Communications Act will apply towards providers of such services.

However, NPT has made it clear that they may, in an interim period, grant temporary exemptions from some of the obligations relating to ECS/PATS providers. This will only be done to a limited extent, subject to individual applications, and under the precondition that consumer interests are adequately protected through marketing information informing customers about potential risks or lacking features.

In its Policy Paper, NPT generally welcomes the nomadic feature of VoIP offerings, although emphasising that the admittance of such use should not jeopardise important end user interests. When it comes to protecting end user interests, the NPT believes that adequate marketing information describing the risks connected with nomadic use can play an important role.

Category 3 VoIP services are the only VoIP services that are currently permitted to use geographical numbers from the Norwegian numbering plan. In addition, a non-geographic 85x series has since 1998 been dedicated to VoIP use. Similar to geographical numbers, the 85x series is currently open only for category 3 VoIP providers.

The use of geographic numbers is contingent upon the service being marketed and appearing as a fixed line telephony substitute, and principally is used from the end-user's permanent address. If the VoIP service is marketed for nomadic use, non-geographic numbers from the 85x series shall be used.

NPT has explicitly stated that porting of geographical numbers from PSTN/ISDN to VoIP providers shall be allowed. The possibility to port numbers is highly appreciated amongst Norwegian end users. Thus, the possibility for end users to keep their old telephone number is expected to stimulate the take up of VoIP services. A user survey carried out by TNS Gallup for NPT confirms this. The survey indicates that 89 % of the VoIP users terminated their PSTN/ISDN subscription when subscribing to VoIP services. Furthermore, 81 % of the VoIP users ported their old fixed line telephone number to the new VoIP service.

Numbers from the Norwegian national numbering plan shall be used in Norway. NPT has therefore stated that the use of numbers from Norwegian number series requires that the provider can demonstrate a relevant nexus with Norway.

⁵⁾ *The Electronic Communications Act implements the EU Regulatory Framework on Electronic Communications.*

In Norway and other countries there has been great interest in how VoIP services may handle emergency calls, or more specifically the obligation to provide correct caller location information to the authorities handling such calls. NPT has stated that, for an interim period, VoIP providers who offer services that can be used nomadically will have an option to be granted temporary exemptions from the emergency calls caller location requirement based on further conditions, *inter alia* an obligation to inform customers about potential risks. However, the call must in any case be forwarded to the emergency response centre, together with the permanent address of the calling subscriber and an indication that this potentially may be a call from a nomadic user.

The reason for giving temporary exemptions is that so far, no adequate technological solution for providing caller location information when the VoIP service is being used nomadically exist. This differs from fixed use of such services, where technological solutions already exist. For this reason, VoIP services that are used on fixed locations only will not be exempted from the obligation to provide caller location information available to authorities handling emergency calls.

NPT has granted about 35 temporary exemptions for VoIP services which may be used nomadically. These are valid until 30 June 2006. NPT is in the process of evaluating whether the time period for these exemptions should be extended.

NPT is also willing to grant exemptions from certain parameters relating to quality measurements specified in ETSI EG 201 769-1. The parameters in question are difficult to apply on VoIP services and give little meaning with regard to quality measurements of such services.

Further, temporary exemptions may be granted from the obligation to offer the calling user the possibility of preventing the presentation of the calling line identification on a per-call basis. Similarly, temporary exemptions may be given from the obligation to offer the called subscriber the possibility of rejecting incoming calls where the presentation of the calling line identification has been prevented by the calling party or subscriber.

Other areas where NPT has indicated its willingness to grant temporary exemptions are the obligation to offer a possibility to block outgoing calls and/or re-

direct calls, and similar, the obligation to provide the possibility to prevent calls being re-directed to the end-user from a third party. Here too, current technical limitations are the underlying reasons why exemptions may be given.

4 Market regulation

NPT is preparing decisions regarding significant market power (SMP) and regulatory remedies in the end user markets for fixed telephony (Markets 1-6) and in the wholesale markets for origination, termination and transit in fixed networks (Markets 8-10). In connection with this work NPT has evaluated whether VoIP services should be included in the relevant markets. NPT has examined a number of factors indicating substitutability between VoIP services in category 3 and PSTN/ISDN, *inter alia* pace of growth of VoIP services, penetration of broadband services, historical evidence of switching in the market, number portability, quality of service and security issues. Especially taking into account the relatively large number of PSTN/ISDN users changing to VoIP, NPT has concluded that VoIP services in category 3 are included in these markets. This conclusion has been supported by the EFTA Surveillance Authority (ESA) in their response to NPT's notification of draft decisions in Markets 8-10.⁶⁾

Telenor, the Norwegian incumbent operator, is found to have SMP in all the end user markets (Markets 1-6). Even though VoIP services are included in these markets, NPT has found that the major competition problems in the markets are related to traditional fixed telephony (PSTN/ISDN) only. PSTN/ISDN is still the dominant technology and constitutes more than 90 % of the connections to the fixed network. Most of the remedies will therefore only relate to Telenor's PSTN/ISDN services, and not to Telenor's VoIP services. For example, Telenor will be obliged to offer a wholesale line rental product for PSTN/ISDN and carrier selection / pre-selection for PSTN/ISDN. These products will not be mandated for Telenor's VoIP services. Likewise, requirements for non-discrimination and transparency will only be mandated for Telenor's PSTN/ISDN end user products, and not for Telenor's VoIP products. As the volume of VoIP services grows at the expense of traditional fixed telephony, NPT expects that the need for specific market regulation of Telenor's PSTN/ISDN services will be further reduced and possibly eliminated.

⁶⁾ NPT's notification of 14 February 2006, ESA's response of 14 March 2006 and NPT's final decision of 24 March 2006 in Markets 8-10 are published on NPT's website.

Regarding the wholesale markets of origination, termination and transit (Market 8-10), the inclusion of VoIP services in category 3 in the markets means, *inter alia*, that VoIP providers who control their own termination and set their own termination charges will have SMP for the termination of calls to their own end users. All providers of termination in fixed networks will be subject to price regulation.

5 Future work

NPT notified draft decisions in Markets 1-6 to ESA in March 2006.⁷⁾ ESA's response is expected in the beginning of April. Depending on the response from ESA, NPT expects to publish the final SMP decisions in these markets in April or early May 2006.

Since VoIP services can easily be provided across borders, differing regulatory approaches in various countries may distort international competition. To

reduce this risk, NPT will in its future work take due notice of European and other international regulatory developments and international market trends.

As mentioned above, NPT is planning to publish a policy paper in the spring of 2006 regarding the status of VoIP offerings that fall under category 1 or 2 in light of the Electronic Communications Act.

NPT will also have an open dialogue with the industry, with the aim of reaching solutions that are of mutual interest to market and regulators. For example, NPT has initiated a joint project related to legal interception control, where also major market players are represented. The aim of the project is to develop technical solutions relating to, *inter alia*, legal interception for VoIP. Affected VoIP providers will be given the opportunity to express their views before the conclusions reached by the group are imposed upon them.

Willy Jensen is General Director of the Norwegian Post and Telecommunications Authority since 2000. As a professor in informatics he has been active in research within computer networking for 30 years. He is presently Counsellor of the International Telecommunication Union and member of the UN WG on Internet Governance.

email: wje@npt.no

⁷⁾ NPT's notification of 2 March 2006 is published on NPT's website.

Business models for broadband telephony

BJØRN ARE DAVIDSEN AND FINN TORE JOHANSEN



Bjørn Are Davidsen works in Business Development at the Fixed Lines Residential Market, Telenor Nordic

The hype and hopes of broadband telephony (BBT) are to a large degree created by the perceived difference in price and business models between the plain old telephony service (POTS) and BBT. This impression is strengthened as long as market and media continue to convey the message that BBT is for free. In this article we look at the reality behind the hype and at the differences in the price models in today's BBT market. One major issue is the difference between the business models for Internet and traditional POTS, in particular related to peering and interconnect termination.

When looking at this area it is important to discover the drivers and the rationale for doing business among various contenders. It is also necessary to distinguish between players who offer BBT as part of their broadband access product, and those who offer BBT as a standalone product that the user may connect to the type of access she already owns.

Related to this there is a difference between selling BBT within the framework of voice business or not. The business model is different if one acts as a traditional telco that looks at BBT as another telephony price offering, compared to having BBT as a value add to another service or product.

A business model needs to consider both infrastructure related costs and service provisioning related costs like customer acquisition, billing and customer service. While the first type of cost may differ in kind between various BBT providers, the second type of cost is similar for every provider with a paid service.

As a general rule it may be said that the overall cost picture is not fundamentally changed when moving from POTS to BBT. The existing rationale for BBT business is a cost-price imbalance. At the same time the price level for Internet services is even less cost related than the price level for telephony. In a perfect market this imbalance will not remain.

To evaluate the impact of BBT we consider the issue of disruptive innovation. BBT bears several characteristics of this.

We also make comparisons between various existing pricing schemes for POTS and BBT. Our conclusion is that the degree to which BBT will succeed in the long run is not due to pricing alone.

Finally we look at the Future of Fixed Phones. Our conclusion is that even if fixed phones in the next ten years will be perceived as a good value proposition for many users, the main game will be migration to broadband mobile offerings.

1 The business of business models

A business model is – not surprisingly – a model for making business. It is about what to sell (the value offering), who to sell it to (the target markets) and what resources are needed (the input costs). Michael Porter very much established the paradigm in this area with his “Five Forces”. In this model one decisive issue was to choose between being low cost and differentiation.

Business models are in general structures that ensure successful commercial practices to take place. Some businesses are quite easy to grasp and describe, others rather difficult. In all cases a lot of thought is needed to develop and structure a viable business model.

The last decades the increasing use of the term “business model” shows a fundamental change in business

thinking. The term symbolises the uncertain role of strategy in today's economy. While strategy had taken a dominant position after Michael Porter in March 1979 (*Harvard Business Review*) published the article “How Competitive Forces Shape Strategy”, it has now moved into more turbulent waters. To Porter, profitability in any industry is determined by five forces: the competition among existing players, the threat of new entrants, the power of suppliers, the power of customers and the availability of substitute products. By analyzing these forces, managers could determine the optimal positioning for their company.

In Porter's world, industries had clear boundaries and stable structures. Success was determined not by the company's quality or innovativeness of its products, but by the logic of its strategy.

It was no coincidence that such a highly codified form of business thinking arrived at the end of the Industrial Age. By the late 1970s, the industrial economy had been going for more than a century. Its structure was fixed and competition predictable. The professional manager had long since replaced the entrepreneur. It has been said that *“The cult of strategy reached its logical, and absurd, conclusion in the 1980s, when managers spent all of their time “restructuring” their companies. Customers, products and employees became unimportant. All that mattered was manipulating assets to earn higher financial returns. Strategy had become an end in itself.”* (Carr 1999)¹.

Carr continues by arguing that in the early stages of an economic system, the rewards go to those who create the new, not to those who conserve the old. Entrepreneurship is more important than stewardship. And since the final form of the new system remains unknowable, strategic “Porter oriented” planning has little use. A new way of thinking about business has gradually emerged. Instead of planning for the future, one looks for business models that reduce the need for planning.

This has facilitated a shift from strategy to models. While a business strategy is a theory – a line of reasoning that ends in a logical conclusion – a business model is a hypothesis. It’s a tentative stab at the truth. This seems to be a good description of the BBT area at the moment. There is an increasing threat from new entrants, customer perceptions are changing and there exist different kinds of substitute products to POTS, from mobile telephony to BBT and PC-based Instant Messaging (IM). This is very much a time of transition and turbulence. It may not be the best of times, or an age of great wisdom, but it is certainly not the worst of times for customers, even if the range of options at the moment is bewildering. Hovering between belief and incredulity, to some it is the spring of hope, not the least for a winter of despair for incumbents.

For this is a time of widely different business models, where similar products are supporting completely different businesses. One is the voice business, with PSTN and ISDN as the old “killer applications”, now being replaced by BBT within the original framework or paradigm that has constituted voice business for the last 130 years. Another is the broadband business, where the rationale behind BBT is driving the growth of broadband. A third model is having BBT as a freebie, usually without interconnect – this is how Skype,

in some way, does business. A fourth is when BBT is set up as a new business, seemingly playing according to similar rules as telephony in the US market, with fixed prices. And against this some have a motivation for consciously disrupting the plain old telephony business, as this is being considered as almost immoral. Having the means, motive and opportunity, new entrants have not been slow in attempting to stab POTS in the back.

2 Comparing business models

A business model can be seen as an answer to the complex question “Who pays what, to whom, and why?” (Clarke 2004)² This definition gives rise to four overall questions that can be used in classification and comparison of BBT business models:

- Who pays?
- What for?
- To Whom?
- Why?

In this article, we will concentrate on consumer services. This most often means that the end-user or someone in the same household or family is the one paying. We will not make a general attempt to describe the vast business world of technology companies or others that take part in delivering components for this end-user offering. But we will see that price and quality differences observed in the consumer marketplace may originate in the underlying industry structure.

The object paid for in the BBT business is generally the ability to talk to other people over distance. The payment can be related to different aspects of this ability, such as fixed monthly service subscriptions and distance based usage charges, calculated per unit of time. If there is no payment involved somewhere, then there is no business model. This does not imply that all models need to be about paying directly for the BBT service as such, as we will come back to.

Payment is traditionally done to a “voice service provider” or to someone operating on its behalf, such as a billing or credit card company. The voice service provider can be the plain old national telco, a local broadband company, or it can be a company operating over the Internet with only virtual presence in the country of the customer.

Customers pay for voice services because they have a need to communicate with other people. However,

¹) “From Strategies to Business Models”, by Nicholas G. Carr, http://www.nicholasgarr.com/articlesmt/archives/from_strategies_to.shtml

²) <http://www.anu.edu.au/people/Roger.Clarke/EC/Bled04.html>

this need in itself is not enough. In order to pay for a voice service, the service should fulfil the underlying desire to communicate better than alternative means of communication. In economic terms, the perceived cost of the service should be lower than alternative costs, as measured by the customer preference, e.g. in terms of money, time, or other resources such as psychological effort. This differentiation with respect to alternative communication means is the core of the value proposal for BBT services. Very few will pay extra for something they already have.

With the perspective of the four answers above, most BBT business models are generally the same as the plain old telephony, or POTS model. It is just a slightly different solution to the same problem for the end user. But this is of course not the whole truth. Depending on the degree of radical thinking, we can oppose to this view in (at least) two different ways:

- 1 BBT is not just slightly different from POTS. It uses the Internet for free and is much cheaper. Therefore it will disrupt the current POTS business.
- 2 The most interesting BBT business models are those that are not covered by the description above. These are the ones that will truly disrupt the whole telecom industry.

2.1 POTS vs. the Internet

Let us first take a closer look at the actual difference between BBT and POTS. Central to this difference is the way the voice signal is transported. This matters relatively little to the average end-user, but is central to the universe of many in the business.

Let us consider a simple case with two users, UA and UB, both with devices (terminals) physically connected to the same network, illustrated as a circle in Figure 1. The total amount paid by UA (thick arrow in figure) shall cover all expenses needed to deliver his voice calls to UB. These costs include investments in network equipment, as well as operating costs connected to the network and to customer care, billing and marketing. In general, the payment can have a fixed *capacity* component, C, and a variable *usage* component U. UB must also pay in order to be connected to the network.

If the network is POTS, the voice service alone must cover all expenses. If the network is IP-based, however, the costs for access and routing are shared by all payable services offered, e.g. an ADSL Internet service, a BBT service and an IP-TV service. The result is that BBT can be offered at a lower rate, although the fundamental network capacity needed to transport voice over IP is generally higher than that needed for

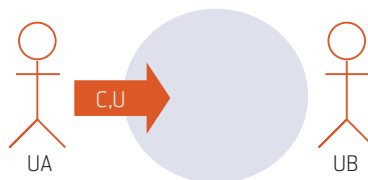


Figure 1 Business model for on-net calls

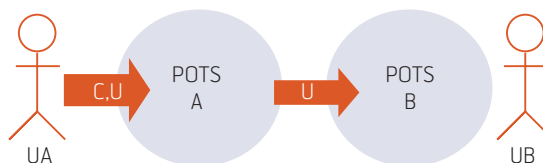


Figure 2 Business model for off-net POTS calls

a modern, digital POTS service, due to large protocol overheads and complex routing logic.

A more complex case arises when the two users are on separate, but *interconnected* networks, A and B, as shown in Figure 2.

In the POTS case, a normal interconnect agreement between operators would mean that network A pays something to network B to terminate calls from UA, as the thinner arrow indicates. Normally, this cost is passed on to the end-user UA, so that the usage price for *off-net* calls is higher than for *on-net* calls. The added price reflects the additional costs of extra traffic in network B, but there can also be a monopoly price component in the amount paid to B, since UA has no alternative way of reaching the subscriber UB. Each network has a *de facto* monopoly on calls to its own customers, and this is one of the reasons why heavy market regulations are still enforced in the telecom business.

The call may be even more expensive for UA if networks A and B have no direct connection between them, so that A has to route the call via a *transit* operator T, as shown in Figure 3.

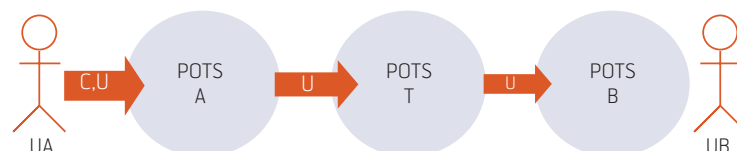


Figure 3 Interconnect business model with transit network

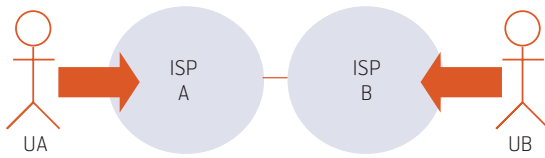


Figure 4 Peering between Internet Service Providers

A transit operator does not need to have end-user customers of its own, but gets a margin from buying local termination from network B and selling it to network A. This margin provides an important incentive for investments in long-distance transport capacity, such as fiber networks, sub-sea cables and satellites. The international transit market is highly competitive, since it can avoid national monopolies to a large extent. This has the interesting side effect that in some countries, calls between competing operators actually flow around the globe before they connect at a transit operator. Incentives for interconnecting locally are not high enough to validate the costs, although the traffic volumes may be high.

The Internet equivalent to POTS interconnect is called *peering*. It works totally different.

As shown in Figure 4, when operators A and B are Internet Service Providers (ISPs), data traffic may flow freely between the networks without any exchange of payment. Both network operators cover their own cost for the traffic and they both keep all the money paid by their respective users. This is known as the *sender keeps all* principle. Network B gets nothing for providing the possible long distance transport from network A to UB, but must recover its cost solely from the amount paid by UB.

However, the Internet also has its own kind of *transit*. Similar to the transit operators of POTS, if two ISPs do not have a direct connection, they can reach each other via a third party, as shown in Figure 5.

The difference between this model and the model in Figure 3, is that money now flows from the end points towards a central place in the network. Small ISPs with only local networks pay *higher-tier* ISPs (like ISP T in the figure) for the ability to reach all other routable parts of the Internet. The transit operators are normally compensated for capacity, not for use, implying that expensive transit connections can become bottlenecks in busy periods.

At the core of the Internet there is a small number of so-called *Tier 1* operators that peer without compensation only with each other. These are the only operators that have the right to route traffic globally without paying for it. All other ISPs must have a commercial transit relation as a last resort routing. All Tier 1 ISPs are based in North America.

One aspect of the peering-and-transit structure is that if two ISPs of different size compete for the same customers in a region, the smallest one will have most to gain from a direct peering relation with the other. Assuming a large amount of traffic is regional only, without a peering relation, the smaller ISP would have higher transit costs per customer than the larger one, which would have more of its traffic on-net. For this reason, a hierarchy of ISPs has developed, and in the future, it may happen that the extent of peering relations will become even less than we see today.

If a voice call is made from two users on different ISP networks, data traffic flows in both directions across the peering exchange point. Only a small

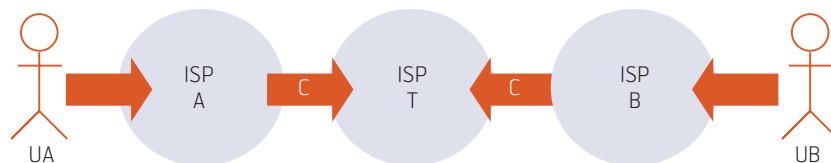


Figure 5 Internet transit via a higher-tier network operator

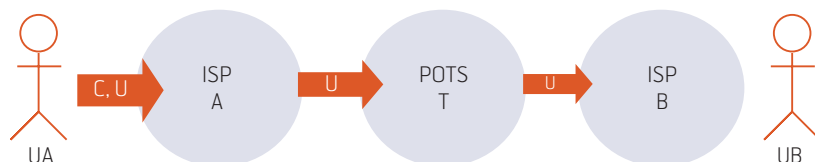


Figure 6 ISPs using POTS transit for voice interconnect

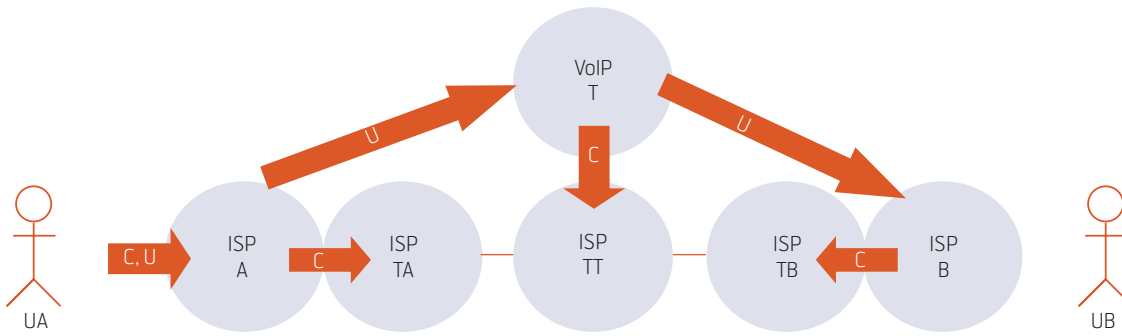


Figure 7 A virtual VoIP transit operator

amount of signalling traffic will tell who of the two parties initiated the call, and this signalling may be hidden to the network operators, for instance by end-to-end encryption. It will therefore not be possible for network A to compensate network B for the traffic generated by the call just by counting bits or IP packets flowing through the peering exchange point. A more complicated arrangement is needed to have interconnect arrangements with compensation models like POTS. For this reason, many providers offering BBT today interconnect via traditional POTS transit operators, as shown in Figure 6.

With the arrangement in Figure 6, quality is regulated by price. Since there is competition among the POTS transit operators, ISP A can choose between different price-quality levels for the transit connection. And since ISP B sells a termination product, which will contribute to the transit operator's portfolio, the price must be at least partly justified by quality measures in ISP B's network. Interestingly, the transit operator POTS T does not need to use traditional technology, only the traditional interconnect business model of buying and selling termination minutes.

At the low end of the price-quality scale, we find the virtual VoIP transit operators, which use low-cost

surplus capacity of other intermediate ISPs to provide call termination.

In Figure 7, the operator VoIP T sells cheap VoIP termination to ISP A. The transit ISP TT provides for the virtual operator's Internet connection and is here assumed to have peering relations with the transit operators ISP TA and TB. Usage-based payment is only done between the end-user ISPs and the VoIP T operator. Thus service and transport are separated. The intermediate ISPs do not see that the traffic generated is voice, and it will receive so-called *best effort* treatment, meaning that other paid-for traffic will have priority in case of network congestions. The number of intermediate ISPs that must be compensated by the VoIP T operator for transit capacity makes this arrangement sensitive to distance, although not in the same way as for POTS.

The principle of separating service and transport can also be applied on a local scale, as shown in Figure 8.

In this case, the virtual operator BBT C competes with ISP A to provide voice services to UA over ISP A's own network. BBT C buys fixed capacity from ISP TT and call termination from one or more operators T, which can provide the best termination rates with acceptable quality. Since the usage payment

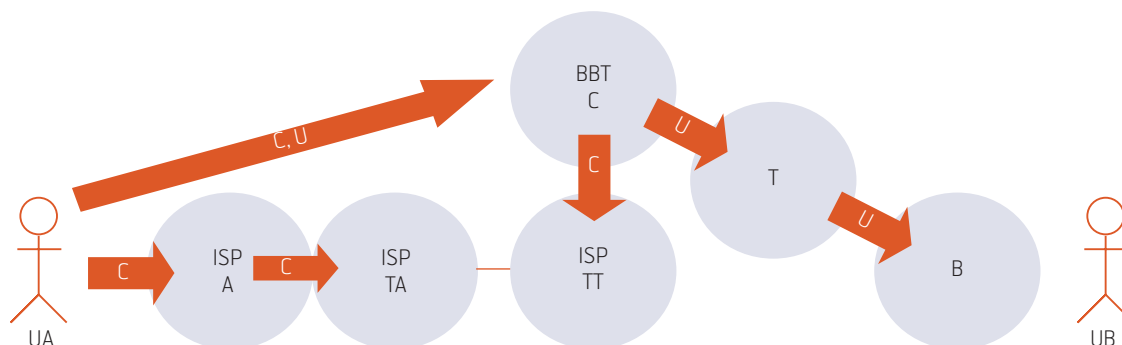


Figure 8 A virtual operator in the end-user market

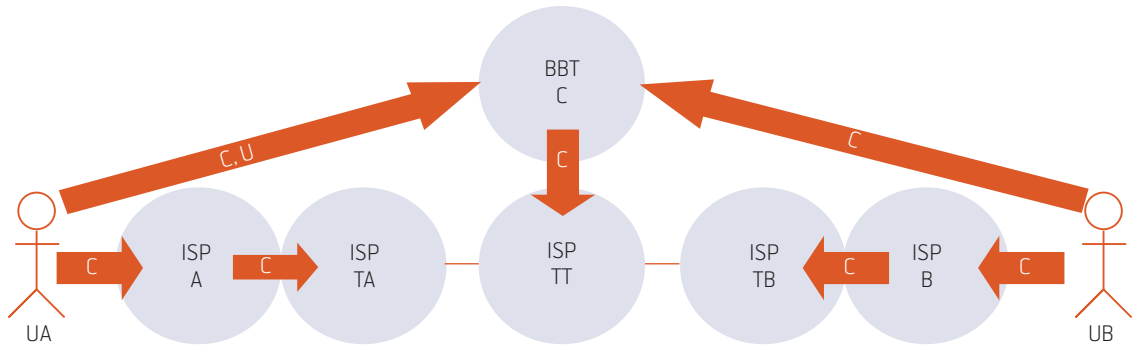


Figure 9 A virtual end-user operator with virtual on-net call

from UA does not follow the call through the network, there are no incentives to provide voice quality in the originating networks of ISP A, ISP TA and ISP TT. Only the terminating part through T and B can be controlled.

A more extreme case is when both UA and UB are customers of the same virtual operator BBT C, as shown in Figure 9.

Here the operator has no control over the end-to-end quality, and is vulnerable to congestion in all IP networks on the route from A to B. On the other hand, there are no termination fees to be paid, and the operator can apply the “sender keeps all” principle to the voice service charges received from UA.

The situation of operator BBT C resembles that of a virtual POTS service provider or *reseller*, who buys *originating* calls from an incumbent operator, on regulated terms, and competes with the incumbent on elements like marketing and customer care, and possibly termination. The important difference lies in the fact that the virtual end-user operator does not have to pay for origination. This means that it does not have the support of the originating network, and will have to rely on the quality of the still unregulated general Internet service offered by ISP A to its customers.

What we have discussed above is that there are actually three different ways of providing more or less the same voice service to the customer:

- Traditional POTS service. The operator has full control over the call, but uses old technology to reach the user.
- Voice as part of a broadband service from the ISP. The operator has full or partial control over the call, and uses new technology.
- Independent broadband telephony service. The operator uses new technology, but has less or no control over the call.

All of these services aim to use the most efficient technology to provide the quality level needed, but the independent BBT service has little control over the network aspect, as the business model provides no incentives for it. The independent model, however, does have a cost advantage in those parts of the call that traverse the Internet. However, the costs of long-distance transport no longer dominate the POTS business, and it is reasonable to believe that both POTS operators and ISPs will be able to compete, since they have better control over the trade-off between quality and price.

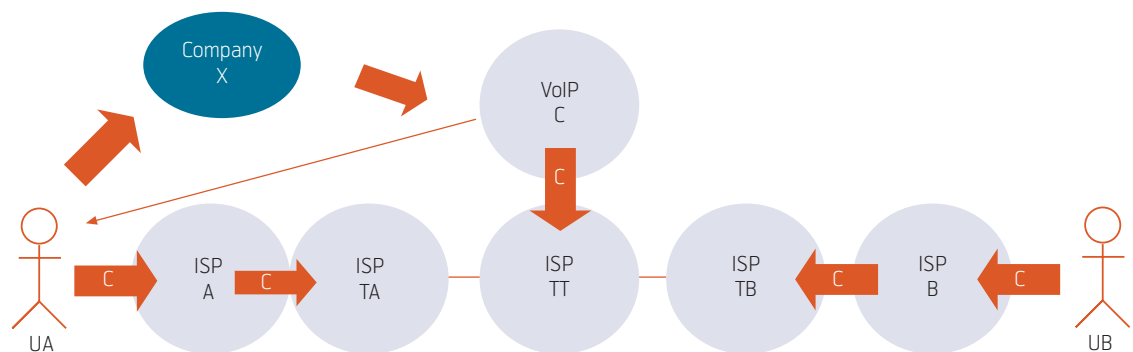


Figure 10 The advertising business model taken to VoIP

2.2 Completely different business models

We should not forget to consider the second line of thought mentioned at the start (which even the most diligent reader has probably already forgotten): What about business models that are *not* similar to POTS? Aren't VoIP and BBT an area of business model innovation and radically different thinking?

So far, the Internet business model that seems to have become most popular is simply *advertising*. Figure 10 shows how this can be taken to BBT, where the user UA receives advertising about company X from the VoIP service provider C, and then receives his VoIP service for free. Company X hopes to regain the marketing expenses from selling its products to UA.

VoIP enables new kinds of advertising – in addition to the now well-known web-click models successfully deployed e.g. by Google. Click-to-call and audio advertisements are two examples. The potential value-creating element in connecting advertising to a communication service is that it can become more targeted. Personal advertisements can be made dependent on the communication pattern, and this will increase the value of an average customer UA for all the companies X financing the VoIP C service.

Of course, these business models have an additional cost of running a voice service, as compared to the normal web sites and search engines with which they compete. The advertising income most probably is not high enough to pay for expensive call termination like mobile and long-distance POTS. Only services with a very large on-net user base will be able to generate profits, and these profits are probably related to more offerings than just voice communication.

A variant of the advertising business model is when company X and VoIP operator C are actually the same company. The VoIP service is then something company X gives away in order to support its main business. This makes sense when the main business has hopes of profit, and when voice expenses are kept low compared to this.

Finally, what currently seems to be a very popular business model for VoIP is the option-based model, as illustrated in Figure 11. The VoIP operator C thinks that if he can just create a large user base through a simple-to-provide, free, or almost free, voice service, he can later think of a way to earn money from these users. Obviously, the valuation of this kind of business model is impossible, but the finance markets have repeatedly shown that just this

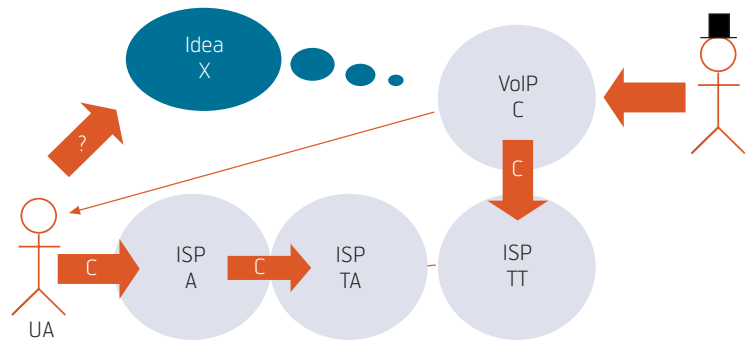


Figure 11 The high-risk business model, also known as lottery

kind of ideas actually receive a reward for being high-risk bets.

To conclude, so far we have probably not seen the survivors in the future business model shoot-out. What we are facing today are free all-you-can-eat voice and even video services, offered to anybody with an Internet connection, without any payment from the end-user whatsoever. This is the reality behind services like MSN Messenger and Skype. The quality is good when the Internet connection is good, the services are user-friendly for anybody sitting in front of the PC, and at the time of writing, Skype does not even bother the user with annoying advertising. With a more suitable terminal, and the ability to call and be reached on ordinary phone numbers, e.g. through an ENUM infrastructure, these services seem like good replacements for home and business fixed phones, if not mobile phones, when wireless Internet becomes ubiquitous.

However, there will probably be even better offers in the future. Mobile phone companies and broadband companies, wireless and fixed, will take the role of Company X and offer the same services bundled with something you still need to pay for; access. With their industry backing, these operators will be able to offer more user-friendly terminals, simpler billing, heavier marketing, as well as services adapted to the local marketplace. In the markets in the rich part of the world, these value offerings will probably be enough to keep the customers, as long as the price is right. And with the same prices for Internet transport available to everybody, there is no reason why it should not.

3 BBT – a disruptive innovation?

Whether one looks at broadband telephony as disruptive because Internet transport is “free”, or because it

3) http://www.readinggroupguides.com/guides/innovators_dilemma.asp

can create fundamentally new business models and value offerings, it is interesting to look at the general theory of disruptive changes. This may lead to answers on how to shape BBT into a successful market offering. The guru of disruptive innovations is Clayton Christensen at Harvard Business School. In an abstract of his book “The Innovator’s Dilemma”³⁾, we find the following rather lengthy, but well-covering definition:

Disruptive technologies change the value proposition in a market. When they first appear, they almost always offer lower performance in terms of the attributes that mainstream customers care about. (...) But disruptive technologies have other attributes that a few fringe (generally new) customers value. They are typically cheaper, smaller, simpler and frequently more convenient to use. Therefore, they open new markets. Further, because with experience and sufficient investment, the developers of disruptive technologies will always improve their products’ performance, they eventually are able to take over the older markets. This is because they are able to deliver sufficient performance on the old attributes, and they add some new ones.

From this definition, voice quality and global connectivity are obvious attributes with POTS that mainstream customers care about. New attributes that mostly apply to new customers may be PC integration, instant messaging, presence and video. However, it is not obvious that BBT actually is a disruptive technology by this definition, at least not for the BBT variants that resemble POTS the most.

In an interview Christensen made with the CIO Magazine April 2001⁴⁾, he elaborates further on a number of litmus tests for successful disruptive innovations:

1 *“in almost every case, a disruptive technology enables a larger population of less skilled people to do things that historically only an expert could do. And to do it in a more convenient setting.”*

This has no relevance to BBT as most users do perceive BBT as “POTS made cheaper”. Viewed from this angle BBT is no disruptive technology. However, if one looks at “calling cheaper” as something only experts were able to do previously, BBT may be a disruptive technology. Conclusion: Not decided in this area.

2 *“The disruptive technology almost always takes root in a very undemanding application, and the estab-*

lished market leaders almost always try to cram the disruption into the established application.”

This has relevance as voice telephony for most users is an undemanding application. This article may also indicate an interest from established market leaders to cram BBT into the existing telephony business model.

3 *“You can’t disrupt a market in which customers are not yet overserved by the prevailing offerings.”*

This has relevance to BBT in several ways. Today’s POTS have more functionality than most people feel a need for, and BBT is perceived as just about cheaper calls. However, there is also the case that mobile voice is perceived as a valid substitute to today’s POTS, due both to pricing and to higher functionality.

4 *“The successful disruptive business model facilitates or lubricates existing patterns of behavior. It’s not predicated on consumers changing behavior”.*

This has relevance to BBT as most users perceive BBT as a form of POTS, and hence something which will not lead to a change in their behaviour. The only difference is in the installation procedure.

5 *“find customers who would be delighted to have even a crappy product because it helps them do what they’re already trying to do better.”*

This has relevance to BBT as it is a crappier product than POTS – and at the moment cheaper.

Based on these five litmus tests, we do conclude that BBT is more a disruptive technology than not. This is also seen in how BBT products are being developed and launched in markets that have unbundled DSL and POTS. While some companies have users with mobile phones only as one of their targets, all companies are targeting POTS customers.

For Telenor there was an assumption that the current POTS products in Norway are overserving at least some customers, since almost everybody has a mobile phone. The Telenor BBT product launched April 2005 was therefore created to serve as a simple home phone replacement for people who don’t want the full reliability, security and quality of a conventional POTS service.

The BBT service itself was designed also to solve fundamental problems for customers that did not have

⁴⁾ <http://www.cio.com/archive/040101/disruption.html>

a home phone, as elaborated in chapter 5 below. The initial Telenor BBT price plan was similar to POTS, only with significantly lower fixed monthly costs. This would minimise the barrier for new customers to buy the service, even though a moderately priced adapter would have to be purchased with the service.

Customers would pay for the service over their broadband bill. This means cost synergies with the broadband offering on marketing, customer care and billing, as well as an economy of scale in the network when more services move to an IP base. The latter is of course part of Telenor's long-term network strategy.

There is little doubt that BBT has a disruptive potential towards the time honoured business model of POTS. Though this does not imply that POTS as a service will go out of business in the near future, the rules of this business are changing.

4 Price propositions – BBT vs. POTS

To most customers the value proposition of BBT is lower price. However, as the examples in Figures 12 and 13 indicate, there is a wide range of price models in the market today. How these translate into business models – not to mention business – is of course another issue, still it does indicate that the range of

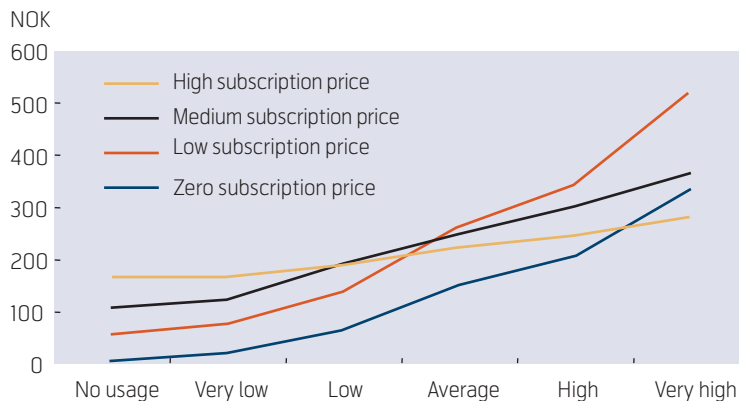


Figure 12 Comparing different BBT offers. Total monthly cost for different usage patterns (Norwegian market, Autumn 2005)

options is higher – and will continue to be higher – than has been the message media has conveyed when talking about BBT for some time.

A “fixed price for everything” proposition like the high subscription fee with no or low usage cost, which has been the dominant in the Norwegian market the last year, is – if one looks closer at it – neither very fixed nor the cheapest for a lot of users. For low to average usage it will be better to have a zero to medium subscription fee and a more normal traffic cost. As the traffic to a higher degree migrates from

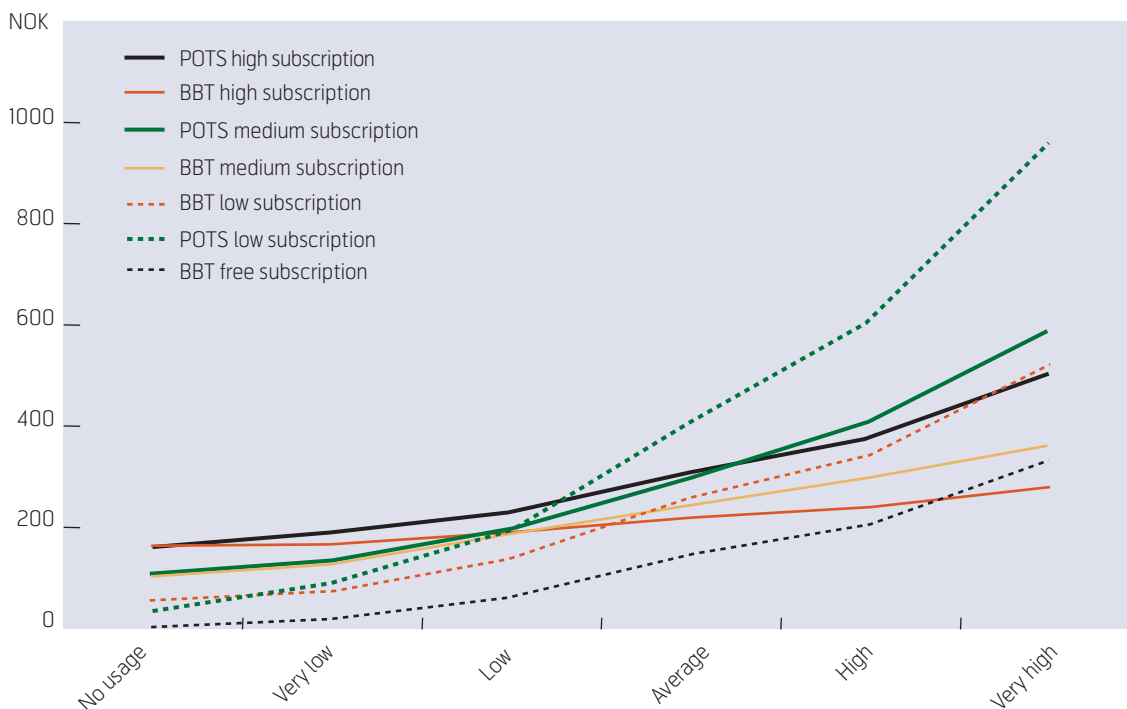


Figure 13 Comparing BBT to POTS offers. Total monthly cost for different usage patterns. Thick lines are POTS⁵⁾ (Norwegian market, Autumn 2005)

⁵⁾ In this figure the line cost (70 NOK) is not included. As some DSL-suppliers do add a fee of about 70 NOK for unbundled DSL, analogue phones are even more favourable for low usage when considering the total monthly costs a customer pays.

	National calls	International calls	Features	Quality	Safety	Other issues
BBT	Somewhat lower prices than POTS for average and high usage	Considerably lower prices than POTS	Fewer than POTS	Lower than POTS, better than mobile Influenced by whether the DSL line also sends high volumes of data upstream during calls	No check of call origin address for emergency calls No emergency power	Positive: • Nomadic use is possible • Gives the user an innovative image Negative: • SPIT may become a threat • Denial of service may become a threat • Uncertain life cycle of terminal adapters (break-downs or outdated), may provide a need for re-buying and a higher cost for BBT than so far perceived
POTS	Perception of too high fixed monthly fees for the most common price models		A broad range of standard supplementary services, few are really in use			
Mobile	Considerably higher than POTS for most calls		More than POTS High speed innovation on terminals (cameras, MP3, radio etc.)	Lower than POTS		Positive: • Mobility • Personalisation • Broadband applications (UMTS) • Hybrid WLAN phones

Table 1 A comparison of price and quality between BBT, POTS and Mobile offerings

POTS to mobile, a fixed price scheme to cover POTS-traffic will be less and less valid.

At the same time however, this is very much about perception. It is psychologically tempting, rather than economically based for a lot of customers. Not the least for those customers who are more concerned with control than cost.

And, importantly, it is possible to find pricing models for analogue telephony that have comparable prices for some types of use, in addition to having better value propositions in areas like quality and safety, at least for the time being.

As Figure 13 shows, for customers with low usage, there exist today POTS pricing plans that in fact are better deals than most BBT price plans.

Overall value propositions are indicated in Table 1.

5 A future for Fixed Phones?

There has always been a need to communicate across distance, whether in war or peace, by beacons or bytes. A talk one-to-one is the most effective way to convey information and emotions in most cases. BBT is based on this long held paradigm.

However, recently a new set of questions has arisen, based on customer needs and new habits. Is there really a need for a fixed “home phone”, POTS or not, in an age of mobility, individuality and personalisation? When considering such issues, it is easy both to miss the forest in favour of the trees and the trees in favour of the forest. On the one hand, there is a tendency among some analysts to look at the *fixed* phone as something doomed to extinction in a near future when mobile phones will cover all needs. On the other hand, there are also voices among both conservatives and radicals, who tend to look at the fixed phone – whether as POTS or BBT, wired or wireless – as something with a long life ahead, as it is sufficiently cheap and convenient for many users.

So, is the future singular? Will mobile conquer all forms of fixed phones? Or are the challenges users meet from having mobile phones only, large enough to create sufficient rationale also for fixed phones?

There are reasons to believe that mobile will not be all, at least not in the near to medium future. There is still – and will continue to be – a marked *price difference* between using a mobile phone and a fixed phone. As new price schemes for POTS and BBT show, the main reason users have had for leaving the fixed services the last years – a high fixed monthly fee – is no longer the only option. Other problems that mobile-only users may experience, like *bad indoor coverage*, *battery running out*, and *low voice quality*, may – taken together – continue to provide sufficient incentives for many users for fixed phone services. Adding to this we have things like *resistance to change* (this will change) – especially among the “plus 40 years” (average age of this group will increase) segment who are used to fixed phone, *fear of mobile handset radiation*, the need to find your *lost handset* (a fixed phone is rather convenient on such occasions), old friends who do not dare to call due to *fear of high prices*, and *safety issues* like address identification when making emergency calls. All in all there will be an experience of increased cost control and “fair pricing” by the introduction of “pay for use” combined with low or no monthly fee services, whether in POTS or BBT, or a “high fee service” with (mostly) free calls proving cost control, in addition to mobile only.

The most common perceived scenario is indicated in Figure 14. Here BBT will grow rapidly the first few years, while POTS declines sharply, and in the end broadband mobile offer will dominate. The rationale behind this view is observations such as

- POTS will have little or no development in services or functionality, even if prices may compete or it may be technologically based on an IP network (POTS as “BBT without a modem or adapter”).
- BBT will be perceived as an innovative product, as well as more often than not cheaper, and users already paying for broadband must need to defend their choice.
- Mobile terminals will be able to access any service on any access.
- The demand for personalisation will lead to mobile phones being ever more important, e.g. for authenticating personal services like mail and TV-channel subscriptions, when looking at hotel and holiday resort TVs.

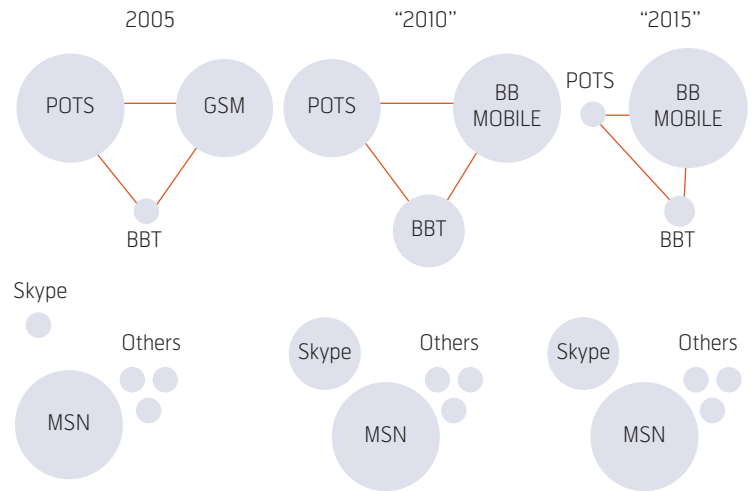


Figure 14 A commonly perceived scenario for the future of voice communication

- Wireless BBT will use WLAN as a carrier and compete with Mobile.
- Mobile (3G) will provide the best overall value proposition (mobility, bandwidth, services).
- Skype, MSN and other Voice over Internet providers will remain closed networks with little or no interconnect to other networks.
- MSN dominates today and will do so in the future, with little or no interest in interconnect.

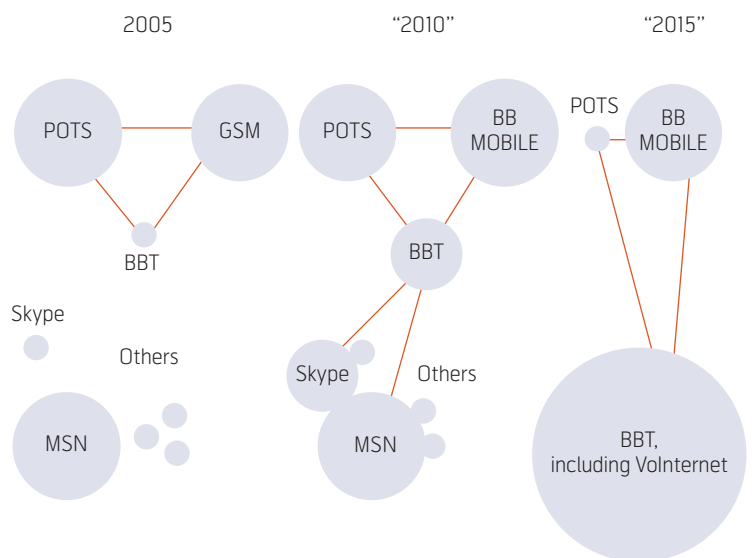


Figure 15 More of a surprise scenario – an enhanced BBT, optimally integrated with WLAN and 3G, dominates due to functionality and convenience, rather than by arbitrating on imbalanced price schemes

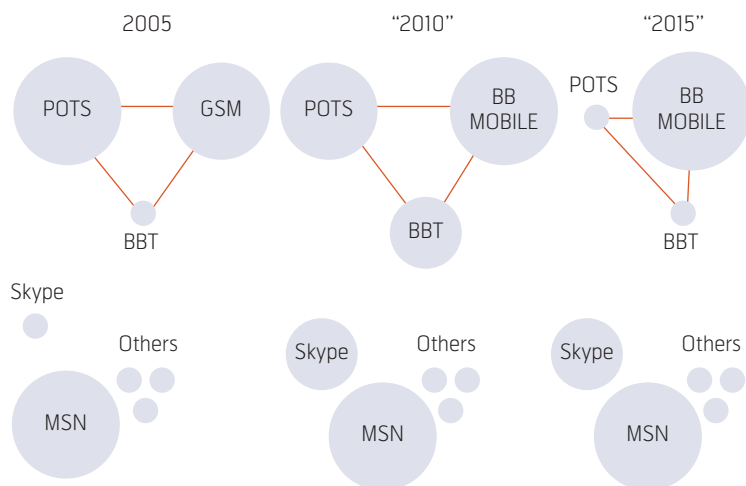


Figure 16 Another possibility is that BBT still will have a lower market share than POTS in ten years, and lower than it will have in five years

A more surprise-oriented scenario is indicated in Figure 15. Here an enhanced BBT, optimally integrated with WLAN and super 3G, dominates. The rationale for this scenario is:

- Mobile (3G) will be too expensive to provide the best overall value proposition.
- Skype, MSN and other Voice over Internet providers become interconnected networks, also with BBT (Skype will continue to grow exponentially and MSN will see the need for interconnect).

An even more unexpected scenario is that BBT will have a lower market share than POTS in ten years, in fact lower than it will have in five years. The rationale for this scenario is:

- BBT succeeds for some years due to pricing arbitrage. However the market for fixed phones will provide a better value and decline slower than in the commonly perceived scenario, while the mobile value proposition continues to improve.
- Wireless BBT will lose out to hybrid BBT/3G services provided by mobile operators.

Which of these scenarios, if any, that will come about does depend on too many factors to model. It will, however, be a surprise if the “surprise free” model of Figure 14 wins out in every detail.

6 Conclusion

BBT does have several of the same characteristics as the Christensen model for disruptive innovation. However, the picture may be more complex than the

commonly perceived scenario of BBT becoming the dominant product in the area of fixed phones.

The drivers and the rationale for doing business are different among various suppliers of BBT. Today some offer BBT as part of their broadband access product, others as a standalone product on any type of broadband access. This leads to a fundamental difference in business models between those marketing BBT as another telephony price based offering, compared to BBT as a value add to another service.

While infrastructure related costs do differ in kind between various types of providers, the cost for customer acquisition, billing and customer service is of a similar type for all. Internet peering models will not necessarily replace traditional voice interconnect, although more voice will eventually be transferred on IP-based networks. As a general rule it may be said that the overall cost picture will not be significantly changed for a voice service provider.

The existing rationale for BBT business is a cost-price imbalance. At the same time the price level for Internet services is even less cost related than the price level for telephony. In a perfect market this imbalance will not remain.

What then, will happen in the future?

BBT is in principle a more future oriented product than POTS, though the production cost will not be very different in the near future. However, even if the future tends to be rather unpredictable, it is not too difficult to foresee that the customer will be the winner. As customers at the moment dream about even more mobility, broadband internet and personalisation, accessed from fancy terminals, to even lower cost, a broadband mobile service, including voice, looks like a rather probable future.

Voice will still be a service that is important enough to pay for; i.e. the demand will remain or even increase slightly. Text messaging and new services like instant messaging and presence are expected to be complements to, and not substitutes for voice communication.

Voice traffic will migrate to the services that can meet the customers' communication demands best. Currently this is the mobile phone. While the present concept of the stationary fixed phone may not be viable in the long or even medium run, a hybrid solution based on mobile as well as WLAN other home based wireless technology may ensure that the bearer will both be the fixed and the mobile network. There may also be a need for a low cost fixed phone for

security reasons. Still, the impetus of habits is hard to overcome. The stationary fixed phone (even without a cordless hand set) may still be with us until today's generations of silver surfers (persons aged above 50 years old) are no more.

As already indicated by some selective POTS offerings, prices for POTS will be reduced to a comparable level with today's BBT prices as the underlying costs of network ownership are reduced.

There will be an increased number of possibilities and options, both related to price schemes, functionality, mobility and personalisation. The overall increased communication value is uncertain so far, though BBT does have a far broader range of options than POTS.

Increased competition, both from broadband access providers and independent services, leads to lower prices in general. Whether telephone companies will do this in a more selective way through BBT or new POTS calling plans, or in a more general way through price reductions for all POTS-customers, remains to be seen.

In a surprise free scenario, the mobile phone will for most customers have the best value proposition for a lot of reasons. At the same time there will be a need for a fixed phone. BBT will grow rapidly and POTS decline even quicker if the POTS operators do not change their pricing model.

The only thing we can be sure of is, however, that surprises will happen.

Bjørn Are Davidsen holds a Master of Science from the Norwegian Institute of Technology (NTH/NTNU), and also courses in Social Anthropology (NTNU), Education (NTNU) and in Master of Management (BI and INSEAD). He has twenty years experience in Telenor in the areas of product development and pilot services within cable television, network development, ISDN and broadband services. The last years Bjørn Are has focused on innovation processes, business development, idea management, creativity and workshop facilitation. Bjørn Are Davidsen has published articles and books on telecommunication, product development, history, science fiction, rock music, cult archeology and science, as well as being frequently asked to speak on such subjects.

email: bjorn-are.davidsen@telenor.com

Finn Tore Johansen holds a Master of Science and a PhD from the Norwegian Institute of Technology (NTH/NTNU). He has 14 years of experience doing scientific research in the field of digital signal processing, speech communication and spoken language technology for telecom applications. He has published various scientific articles and conference papers in these areas. For the last six years, Finn Tore has been involved in business development and strategy projects for person-to-person communication services within Telenor. He also enjoys scuba diving and a good vintage port.

email: finn-tore.johansen@telenor.com

Telephone Number Mapping (ENUM) – A short overview

TROND ULSETH



Trond Ulseth is Senior Research Scientist at Telenor R&D

This article presents a brief overview of ENUM, a mechanism that allows the translation of traditional telephone numbers into a format that can be used to store and retrieve Internet addressing information.

Introduction

The traditional circuit-switched telephone network uses an addressing mechanism defined in ITU-T Recommendation E.164 [1]. The addressing mechanism in an IP network is often referred to as an URI [2]. These are two different mechanisms, and may exist independent of each other. The addressing of a VoIP subscriber may without any problem use the E.164 number in the format sip://1234567890@telenor.no. However, a large-scale VoIP system requires a system where the telephone numbers map onto an IP addressing mechanism – Telephone Number Mapping.

What is ENUM?

ENUM is a solution to the question of how network elements can find services on the Internet using only a telephone number, and how telephones, which have an input mechanism limited to twelve keys on a keypad, can be used to access Internet services. ENUM has been developed by IETF WG ENUM, and is defined by IETF in RFC 3761 [3]¹⁾. ENUM is also adopted by ITU-T and ETSI.

ENUM is built on top of DNS, and specifies transformation of E.164 numbers into DNS names enabling the use of existing DNS services. ENUM is only applicable for E.164 numbers.

The domain “e164.arpa” is being populated in order to provide the infrastructure in DNS for storage of

E.164 numbers. In order to facilitate distributed operations, this domain is divided into subdomains as illustrated in Figure 1.

Tier 0 is administered by IANA in cooperation with ITU-T and the national regulatory authorities.

The administration of Tier 1 and subdomains below is a national responsibility. The Tier 1 domain name reflects the E.164 country code. The Norwegian Tier 1 domain name is therefore 7.4.e164.arpa. (The E.164 country code for Norway is 47.) For Europe the administrative requirements are specified in ETSI TS 102 051 [4].

The number may then be registered for one or more ENUM services. For example, a subscriber may wish to register a telephone number to receive calls at a home phone, or at the office. Additionally, that subscriber may wish to register an email address, as well as a fax machine, to match the telephone number.

It should be emphasized that ENUM is not a normal DNS registration mechanism; it is a conversion of a number already allocated (E.164) according to rules established by ITU-T.

Why ENUM?

The simple answer to the question is already given above; ENUM provides a mechanism for how network elements can find services on the Internet using only a telephone number, and how telephones, which have an input mechanism limited to twelve keys on a keypad, can be used to access Internet services. The obvious rationale for ENUM is the introduction of voice services in an IP network, but as already described more services may be associated with a single telephone number.

How ENUM works

As already stated ENUM is a method to convert a regular telephone number (e.g. +47 85 05 09 62) into a format that can be used on the Internet within the Domain Name System (DNS) to look up Internet

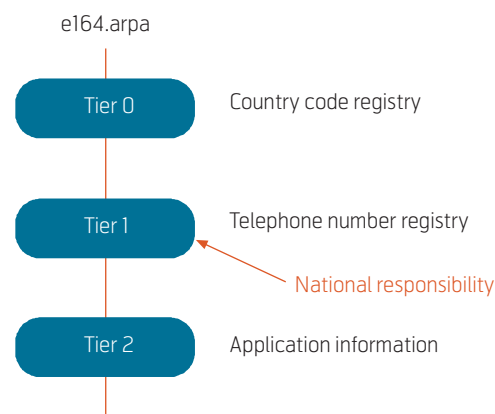


Figure 1 ENUM domain structure

¹⁾ IETF RFC 3761 is the second version of the protocol. The first version was defined in RFC 2916.

addressing information such as Uniform Resource Identifiers (URIs). In the regular telephone system, the most significant number appears first, for example the country code +47 for Norway (the '+' character is a substitute for the national prefix; usually '00'). In Internet domain names the most significant information appears last – for example www.telenor.no. The country information 'no' is last, but will be the first resolved to find the top-level domain for Norway. The principles for this conversion are described in Figure 2.

First, all non-digit characters and/or national prefix are removed. Then the digits are arranged in reverse order and dots are inserted between each digit. Finally, the domain name e164.arpa is added at the end.

This information is stored within the domain name system (DNS), providing routing information to reach the device with the associated ENUM number.

Another feature of the ENUM protocol is that more than one contact information can be stored in the DNS record that is belonging to a specific ENUM number. An ENUM record might contain instructions for a VoIP call (e.g. h323: telenor-operator@telenor.no or sip: telenor-operator@telenor.no), a facsimile call (e.g. fax: telenor-fax@telenor.no), e-mail communications (e.g. mailto:firmapost@telenor.no). Additional services can be developed in the future to be included in the ENUM name records.

This facility would allow the phone number in ENUM to be the single contact number for multiple contact methods for any type of communication (voice, fax, e-mail, mobile, text messaging, location based services, web pages [3]).

RFC 3761 requires that ENUM services are defined in an RFC and officially registered with IANA, the organisation responsible for assigning IP addresses and IP protocol codepoints. The following ENUM services are registered so far:

- h.323 defined in RFC 3762 [5]
- sip defined in RFC 3764 [6]
- presence defined in RFC 3953 [7]
- web and ftp defined in RFC 4002 [8]
- email, fax, sms, ems and mms defined in RFC 4355 [9]
- voice defined in RFC 4415 [10].

Infrastructure ENUM

Traditional (Public) ENUM is a capability provided for end users and is optional for both the calling and

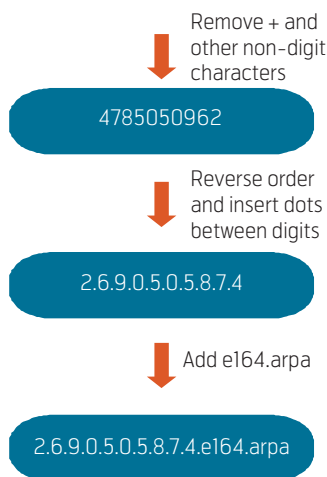


Figure 2 ENUM conversion principles

the called user. The individual user decides whether to register in an ENUM registry.

Up till now ENUM according to RFC3761 (User ENUM) has not been seen as useful to an NGN/ Telco/VoIP provider as it depends on user action in terms of registration, insertion of data and management of that data. Service providers and/or network operators cannot base their services on an optional technology outside their control.

Recently, another type of ENUM called “Infrastructure” ENUM has been proposed. The terms “Carrier” ENUM or “Operator” ENUM have also been used to identify this functionality. The ETSI Technical Report TR 102 055 [11] clarifies the term, indicating possible evolution paths and describing usage scenarios.

The basic principle of “Infrastructure” ENUM is to provide information only to IP communication service providers; some providers may even want to provide this information only to selected peers. The end user has either no access to this information, or he may not be able to use it. This purpose is incompatible with the principle described in RFC 3761 [3], because “Infrastructure” ENUM needs the full population of the information at least for the number range in question. Hence, it must be implemented as an independent system.

“Infrastructure” ENUM technology may also be used to provide access to national number portability information stored currently in IN databases. The problem with this information is that it has only national significance, for example national routing numbers. This kind of data can therefore not be used directly in supranational Infrastructure ENUM implementations.

These issues need to be solved. Currently there are a lot of discussions on the IETF ENUM reflector.

ENUM trials and available services

Shortly after the approval of the first version of the ENUM standard in 2000 both ITU-T and national regulators begun considering ENUM. Workshops were held, and soon ENUM trials were set up in several countries. Among the countries carrying out ENUM trials are

- USA
- Several European countries (Including France, Germany and UK)
- Japan
- Australia.

A few countries have already established a public ENUM service. The first service was offered in Austria, starting in December 2004 [12].

Conclusion

ENUM will be a key driver for the increasing convergence between IP based networks and networks offering telephony service such as PSTN, ISDN and GSM. The possibility to associate a single E.164 number with a list of URIs allows an end user to have a single contact point (E.164 number) corresponding to a number of different services and applications such as voice, e-mail, fax, unified messaging, etc. The end user, by using the functionalities provided by ENUM, can customize his service profile and determine the preferred way to be contacted by the party initiating the communication. ENUM will therefore be beneficial for the users and provide new business opportunities for service providers.

References

- 1 ITU-T. *The International Public Telecommunication Number Plan*. Geneva, 1997. (ITU-T Recommendation E.164)
- 2 IETF. *Uniform Resource Identifiers (URI): Generic Syntax*. 1998. (RFC 2396).
- 3 IETF. *The E.164 to Uniform Resource Identifiers (URI) Dynamic Delegation Discovery System (DDDS) Application (ENUM)*. 2004. (RFC 3761)
- 4 ETSI. *ENUM Administration in Europe*. Sophia Antipolis, 2005. (ETSI TS 102 051)
- 5 IETF. *Telephone Number Mapping (ENUM) Service Registration for H.323*. 2004. (RFC 3762)
- 6 IETF. *enumservice registration for Session Initiation Protocol (SIP) Addresses-of-Record*. 2004. (RFC 3764)
- 7 IETF. *Telephone Number Mapping (ENUM) Service Registration for Presence Services*. 2005. (RFC 3953)
- 8 IETF. *IANA Registration for Enumservice 'web' and 'ft'*. 2005. (RFC 4002)
- 9 IETF. *IANA Registration for the Enumservices email, fax, sms, ems and mms*. 2006 (RFC 4355)
- 10 IETF. *IANA Registration for the Enumservice Voice*. 2006 (RFC 4415)
- 11 ETSI. *Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); ENUM scenarios for user and infrastructure ENUM*. Sophia Antipolis, 2005. (ETSI TR 102 055 v1.1.1)
- 12 *Ipcom GmbH*. [online] <http://www.enum.at/>

For a presentation of the author, please turn to page 2.

Multimedia over IP Networks

ANDREW PERKIS, PETER SVENSSON, ODD INGE HILLESTAD, STIAN JOHANSEN, JIJUN ZHANG, ASBJØRN SÆBØ AND OLA JETLUND



Andrew Perkis is Professor at The Norwegian University of Science and Technology, Trondheim

Multimedia communication over IP networks can either be one-way from a sender to one or more receivers (e.g. streaming) or two-way or interactive multi-way between two or more communicating parties (e.g. video or audio conferencing). Both these scenarios rely on audiovisual coding systems and network protocols such as TCP, UDP and RTP. The coding systems introduce coding distortions while the networks introduce loss and/or delay of the information resulting in distortions in the decoded signal. This distortion needs to be quantified in order to further measure the perceived quality of the media presentation. These measurements can then be used in a feedback loop through the network to adapt the coding systems in order to enhance the user's perceived quality. This paper gives an introduction to the field, identifying the major challenges lying ahead of us and some of the solutions chosen to solve them. The case studies investigated are streaming media and audiovisual conferencing.



Peter Svensson is Professor at The Norwegian University of Science and Technology, Trondheim

Introduction

The simplicity and flexibility of packet switched communication using the IP protocol has played an important role in its emergence as the method of choice for multimedia delivery. However, multimedia over IP and wireless networks faces many challenges due to network variability and lack of service guarantees with respect to available bandwidth delay and jitter. These result in packet-loss or in packets being delivered after they are required. Clearly, the effect of such losses on video depends on how the video stream has been coded and how it has been mapped into IP packets.

In the current best-effort Internet service model, no service guarantees with respect to packet loss, delay jitter and available bandwidth can be made. Packet loss most often occurs due to congestion in network nodes; more and more packets are dropped by routers in IP networks when congestion increases. While packet loss is one of the things that make the TCP protocol efficient and fair for non-real time applications communicating over IP networks, the effect of packet loss is a major issue for real-time applications such as streaming of audiovisual media using the RTP protocol over UDP/IP. Even delay jitter manifests itself as packet loss, since packets received after the intended playout/presentation times are not useful.

The Centre for Quantifiable Quality of Service in Communication Systems – Q2S – deals with Quality of Service (QoS) issues in heterogeneous, multilayered networks where packet switching technology is employed. By services is meant traditional tele-services along with multimedia, messaging, web and information services, as well as location and content

aware services. The Centre works within the following areas: dependability, traffic and security as applied to multiparty communication, as well as two research areas relevant to this paper: Audio over IP networks and Multimedia over IP networks.

Currently, providers and authors of multimedia presentations have to create multiple formats of content and deploy them in a multitude of networks in order to meet consumers' increasing demand for high quality interactive multimedia. Also, no satisfactory automated configurable way of delivering and consuming content exists that scales automatically to different terminal and network characteristics, device profiles or QoS. The quest to represent, deliver and present such interactive multimedia with the ultimate experience in multimedia entertainment and conferencing in such a multimedia framework means that the boundaries between the delivery of audio (music and speech), accompanying artwork (graphics), text (lyrics), video (visual) and synthetic spaces will become increasingly blurred.

In Section 2 we give an overview of audio visual coding, focusing on new video coding techniques including H.264/AVC (H.264 named by ITU and MPEG-4 Part 10 Advanced Video Coding named by ISO/MPEG) and scalable coding. As usage examples Section 3 describes the major challenges in streaming media, while Section 4 targets audio conferencing. Section 5 deals with perception, specifically quality as perceived by a receiver/listener or by the communicating parties, of the resulting sound fields and of the audio based communication process. For video over IP networks we describe two new metrics based on blockiness and packet loss.



Odd Inge Hillestad is a PhD student at Q2S at The Norwegian University of Science and Technology, Trondheim



Stian Johansen is a PhD student at Q2S at The Norwegian University of Science and Technology, Trondheim



Lijun Zhang is a PhD student at The Norwegian University of Science and Technology, Trondheim

Audio visual coding

Possibilities to compress speech, audio, images and video have been explored ever since digital techniques were introduced. There has been a strong emphasis to minimize the bit-rates of such media content in order to use limited bandwidth and/or storage space resources efficiently for rapidly increasing communication demands. The strive for efficient compression algorithms has for some applications, such as two-way communication, been balanced by the requirement to keep processing delays low. Another central factor is the robustness to errors. For speech, a number of codecs (coder/decoder) have been developed and used in e.g., mobile telephone systems, ranging from simple logarithmic quantization to advanced speech modeling codecs in use today that are based on predictive coding. In audio, the compact disc was the breakthrough for digital audio using uncompressed audio for maximum quality with mechanisms for a high robustness to drop-out errors. In the 1990s compression of audio started to be used in applications like digital radio broadcasting and digital audio distribution formats such as Sony's Minidisc. Compressed audio got a real breakthrough in the form of music file transferring on the internet. The well-known MP3 format which is one example of ISO/MPEG coding [MPEG-1 1991, MPEG-2 1994] has led to a wide acceptance of such compression. A rapid development has been possible because of formats such as Real Audio and Windows Media Player. Consequently, the bit-rate needed for what is judged to be adequate quality has been decreasing steadily.



Asbjørn Sæbø is an audio consultant currently employed as a post.doc. at Q2S at The Norwegian University of Science and Technology, Trondheim



Ola Jetlund is a postdoc at Q2S at The Norwegian University of Science and Technology, Trondheim

As of today most of the successful techniques for video coding are in some way a part of either the MPEG standards or the Society of Motion Picture and Television Engineers (SMPTE) standards. SMPTE is currently in the process of standardizing a low rate video codec referred to as VC1, which again is based on the Windows Media Video (WMV) codec developed by Microsoft. Vendor technologies are also included in the MPEG standards. The most current of these is the MPEG-4 standard. There also exists one major open source video codec that is based upon MPEG-4, namely OpenDivX. Video coding or more specifically video *compression* is in the literature separated into four categories [Ebrahimi 1998]; waveform, object based, model based, and fractal coding. The first is also commonly referred to as transform coding. The first three of these categories are included in the MPEG standards. All in all, the MPEG standards comprise a large knowledge base of video coding techniques and, because of the ongoing efforts by the Moving Pictures Experts Group provide state-of-the art video coding techniques.

MPEG-1 covers video compression with target bit rates up to 1.5 Mb/s and consists of techniques for synchronization and multiplexing of audio and video, compression of non-interlaced video, and a compression codec for audio designed for perceptual coding. The MPEG-2 standard is typically used for encoding audio and video broadcast signals for target bit rates between 1.5 and 35 Mb/s. Unlike MPEG-1 this standard also includes interlaced video. Enhanced versions of MPEG-2 are used as the video codec in DVD movies and in most HDTV transmission systems. The MPEG-4 standard includes many of the features found in MPEG-1 and MPEG-2. However, this standard also includes support for digital rights management, 3D-rendering, and target bit rates as low as 8 kb/s and up to 35 Mb/s. MPEG-4 consists of numerous parts dealing with system description for synchronization and multiplexing, but also with state-of-the-art coding techniques for audio and video such as; Advanced Audio Coding (AAC), Advanced Video Coding (AVC) and carriage on IP networks.

The newest video coding standard is a joint effort between ISO and ITU resulting in many names. The standard is commonly referred to as either MPEG-4 Part 10 Advanced Video Coding or H.264. For simplicity we will refer to the standard as H.264/AVC. H.264/AVC has achieved a significant improvement in compression performance, error resilience and a "network-friendly" video representation.

One major challenge in video compression is the transmission of video in lossy environments. A solution is to make packets transmitted in real-time multimedia environments self-contained [Tamhankar 2003]. Thus, no packets rely on other packets in the reconstruction process. As an example H.264/AVC defines a network abstraction layer (NAL), in addition to the video coding layer (VLC) that allows for using the same "video syntax" in multiple environments.

[Wiegand 2003, Stockhammer 2003] gives a good overview of H.264/AVC focusing on the video coding layer (VCL) and the network adaptation layer (NAL) as shown in Figure 1. It has been shown that the NAL design specified in the standard is appropriate for the adaptation of H.264 over RTP/UDP/IP [Wiegand 2003].

The VCL consists of the core compression engine and performs all the classic signal processing tasks. It is designed to be as network independent as possible. The VCL comprises syntactical levels known as the block, macro block, and slice level. The VCL contains coding tools that enhance the error resilience of the compressed video stream.

The NAL defines the interface between the video codec itself and the outside world and adapts the bit strings generated by the VCL to various network and multiplex environments in a network friendly way. It covers all syntactical levels above the slice level and operates on NAL units, which give support for the packet-based approach of most existing networks.

A NAL unit (NALU) is effectively a packet that contains an integer number of bytes. The first byte of each NAL unit is a header byte that contains an indication of the type of data in the NAL unit, and the remaining bytes contain payload data of the type indicated by the header. The payload data in the NAL unit is interleaved as necessary with *emulation prevention* bytes, which are bytes inserted with a specific value to prevent a particular pattern of data called a *start code prefix* from being accidentally generated inside the payload. The NALU structure definition specifies a generic format for use in both packet-oriented system and bitstream-oriented transport system, and a series of NALUs generated by an encoder is referred to as a NAL unit stream.

Currently, three major applications for H.264/AVC may be identified by using the IP protocol as a transport [Wiegand 2003]:

- The download of complete, pre-coded video streams. Here, the bit string is transmitted as a whole, using reliable protocols such as ftp or http. There are no restrictions in terms of delay, real time encoding/decoding process and error resilience.
- IP-based streaming. In general, it allows the start of video playback before the whole video bit stream has been transmitted, with an initial delay of only a few seconds, in a near real-time fashion. The video stream may be either pre-recorded or a live session, in which the video stream is compressed in real-time, often with different bit rates.
- Conversational applications, such as videoconferencing and video-telephony. For such applications delay constraint apply significantly less than one second end-to-end latency and less than 150 ms as the goal so real time encoding and decoding processes are main issues.

In addition, the use of H.264/AVC coded video in wireless environments is described in [Stockhammer 2003].

The development of efficient scalable coding schemes is motivated mainly by the possibility of adapting the encoded data to match channel/network

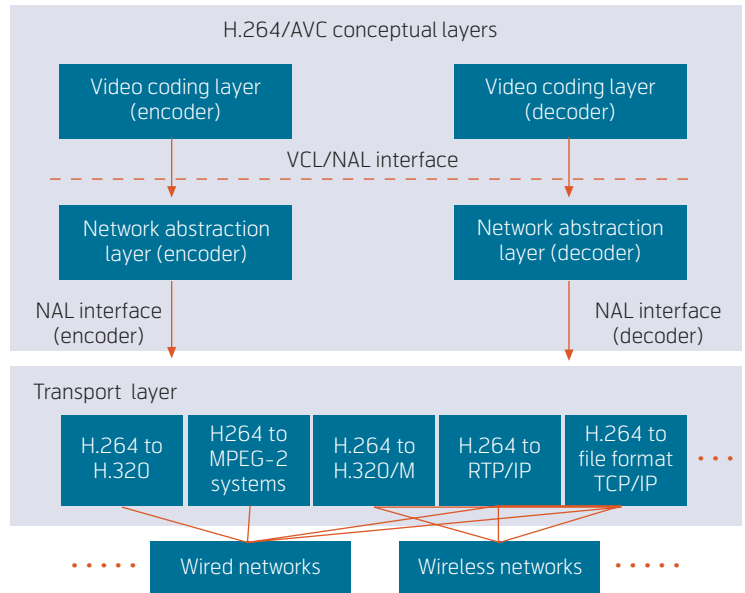


Figure 1 Transport environment of H.264/AVC

conditions, terminal capabilities or business models. For video, scalable coding was included in MPEG-4 (fine-granular scalability, FGS) [Radha2001]. A scalable extension of the H.264 standard has been proposed. Both of the two coding schemes are the traditional block-based *hybrid* type; that is, the coding is done by first identifying and compensating for motion between successive frames and then encoding the residual image using a still-image-like coder. Such coding schemes are not ideal for scalability, due to the so-called “drift” problem (encoder and decoder mismatch). Since the encoder uses the full-rate, full-resolution encoded frames as references for its inter frame prediction, an asymmetry occurs between encoder and decoder when decoding a lower-rate or lower-resolution version of the resource. Due to this, more inherently scalable video coding schemes have been investigated. 3D sub band video coding uses, as the name suggests, a sub band decomposition in both time and space prior to quantization and subsequent processing. This can easily be combined with embedded coding schemes, resulting in an embedded bit stream that, by definition, is scalable (decoding can be stopped at any time in the bit stream). This also eliminates the drift problem. Recent results [Ohm2004, Chen2004] indicate that performance comparable to (non-scalable) H.264 can be obtained with these 3D coding schemes for most sequences. Furthermore, since the generated bit stream is embedded, effective unequal error protection (UEP) schemes can easily be applied. These schemes utilize the fact that the importance of each bit is (conceptually) monotonically decreasing as one moves from the start to the end of the bit stream. Different parts of

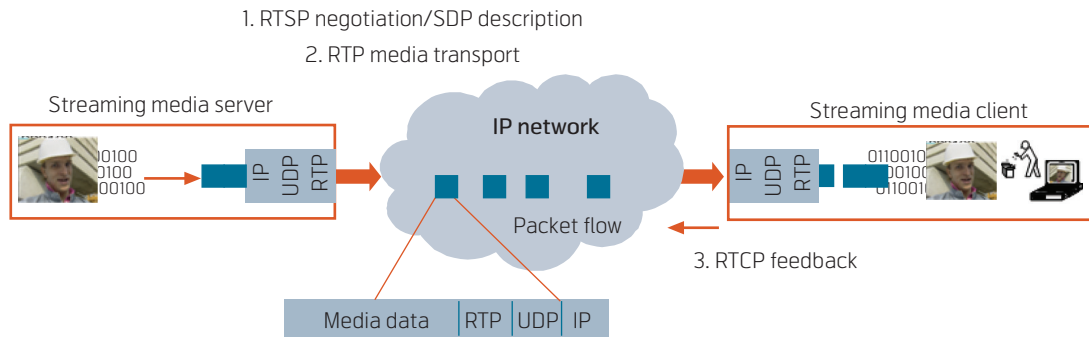


Figure 2 Streaming media architecture

the bitstream are assigned different strength error correction codes, dependent upon the relative importance of the bits and the channel/network conditions. Novel results include [Albanese1996] [Mohr1999] [Puri1999], and fast algorithms for assigning UEP have been proposed [Stankovic2002]. Recently, this concept has been extended to transmission over multiple parallel channels with possibly different characteristics [Johansen2005].

Applications of audio visual communications

Many important challenges arise when considering audio visual communication over packet-switched networks such as multi service IP networks. Available resources in these networks are shared among competing flows with highly varying characteristics, and even if the network supports some QoS scheme (like e.g. service differentiation) individual media flows will typically experience varying delay jitter and occasional packet loss during periods of network congestion.

To prevent packet loss from corrupting an entire file or session, application protocols like FTP or HTTP employ TCP, a reliable byte-oriented transport protocol which uses acknowledgements to explicitly state which byte segments that have been properly received. Lost segments are automatically retransmitted. While this mechanism obviously does not work well for transport of media with real-time constraints, the unreliable real-time alternative, RTP over UDP, might employ a similar approach on the application level for retransmitting lost packets. Typically, such an automatic repeat request (ARQ) mechanism would only be used for retransmitting packets belonging to frames that could still arrive in time for decoding and presentation.

Streaming media

For an audio visual streaming system, the codecs being used and their configuration are of fundamental

importance, and decide the upper limit for the reconstructed quality and level of distortion introduced. Depending on the application scenario (e.g. live broadcast, two-way conversational or on-demand streaming), additional encoder decisions with regard to rate control and coding mode, error control and error resilience tools (e.g. Forward Error Correction, FEC, data partitioning), rate shaping and packetization all have an important effect on the audiovisual quality and application performance during a video streaming session. [Wu01]

Furthermore, the perceived quality heavily depends on the system's ability to adapt to changing network conditions and to show a graceful behavior in case of lost or delayed packets. For instance, packets that are delayed beyond the time they should be available for decoding at the receiving end, are of no use and inflict the same damage as loss of packets in congested router queues. This situation can be alleviated by increasing the size of a playout buffer (also known as jitter buffer) in the streaming media client, but this increases the end-to-end delay and memory requirements. Another recent proposal to prevent buffer underflow is adaptive media playout, in which the playout rate of decoded frames is altered in accordance with the current buffer fullness [Kalman04].

Intelligent retransmission schemes can also be used to recover lost packets and help prevent loss of media data, but in situations of heavy network congestion packet loss is destined to occur. In case packet loss indeed occur, the aspect of graceful behavior is reflected in the decoder's error resilient behavior; the use of proper error detection and error concealment techniques makes sure that the visual appearance of the media resource shows a gracefully degrading visual reconstruction. [Wang98]

Figure 2 depicts a typical streaming session, in which a piece of media content is delivered from a server to a client on-demand. The client requests the media using RTSP, and receives a description of how to

access and decode the media flows for this session. The media are transported using RTP/UDP, and RTCP may be used to feed reception statistics back to the server. [RTP96] At the receiving end, the streaming media client retrieves media packets from the playout buffer, detects packet loss, decodes available media data and tries to conceal missing parts of the media representation. To measure media quality on the receiver side, and help make the right decisions in adapting the delivery, objective quality metrics could be applied to the reconstructed media in order to estimate end-user perceived quality.

At Q2S, a dedicated multimedia test bed and IP networks are used to measure the performance and characteristics of streaming media applications. Besides a streaming server and a streaming media client, it consists of an IP network emulator that enables us to subject our application to varying delay and packet loss in real-time. In addition, the test bed is equipped with high-performance packet monitoring and capture devices, and a packet flow regenerator that enables us to replay media streams and recreate specific scenarios and network conditions. [Hillestad05]

Audio conferencing

Audio conferencing, either standalone or in combination with video, spans a wide range of existing and possible applications. These applications may be grouped into classes dependent upon their quality, bandwidth, latency and other requirements. Typically, the low and medium requirement applications are already in common use, while more advanced and demanding applications are at a research stage.

IP telephony (“Voice over IP”) is in rapid growth, and is an example of a moderate requirement application. Audio bandwidth is low (typically 3 kHz), and either coarse resolution (logarithmic PCM, G.711) or efficient coding (e.g. G.723, G.729, iLBC) is applied. A single channel is transmitted, at a low bitrate (from 64 kb/s to a few kb/s). This is a two-way service, but relatively high latencies (up to 150 ms) are accepted [ITU03].

Medium requirement applications may be exemplified by video conferencing. There is at least one speech bandwidth audio stream and one video stream. Additional audio streams may be speech bandwidth, but may also be full-bandwidth streams for transmission of “CD-quality” audio. There may also be additional video streams. Audio streams are encoded with speech or audio coding. Latency requirements are as for telephony. Bit rates may range from 128 kb/s (typ. lower limit for high quality audio) and upwards to several Mb/s.

The highest requirements may be found for telepresence applications, where a distributed shared multi-sensory immersive environment is formed. In [Woszczyk05] a system capable of transmitting digital video, 24 channels of audio and 4 channels of vibration is reported. For video, SDI at 270 Mb/s is used, with a transition to HD-SDI at 1.5 Gb/s being planned. Audio channels may be 2.3 Mb/s each (24 bit @ 96 kHz).

This last conferencing example hints at possibilities for one way streaming (capture, transmission and re-creation) of realistic sound fields. This will typically be a high requirement application. The number of channels is dependent upon the technology chosen. Binaural technology utilizes two channels [Sæbø01, chapter 2.3] or up to two channels per listener. Ambisonics demands at least three or four channels, while Wave Field Synthesis and other formats may require a very large number of channels [Daniel03]. In all cases, these channels should be full-bandwidth and of high quality, with “CD-quality” (16 bit @ 44.1 kHz PCM, giving 705 kb/s per channel) as a minimum. Professional production quality (24 bit @ 96 kHz PCM, 2.3 Mb/s) may be required. Latency requirements may not be very strict. Where latency and quality requirements permit, it may be feasible to encode the signals to reduce the data rate. The audio codec AAC is considered as basically giving transparent sound at 128 kb/s.

At Q2S, work on audio conferencing has mainly been focused on distributed music playing applications. These networked joint musical performances consist of distributed participants, connected by an IP network, playing together. As a “light” telepresence application, requirements are strict. Audio should be full audio bandwidth, of high quality, and probably multi-channel. Transmission of one or more synchronized video streams would be a valuable addition. For the sake of synchronization between the musicians, latency should preferably be kept below 20–40 ms. Work at Q2S has comprised experimental work and implementations, and a new tool for such applications (LDAS – Low Delay Audio Streamer) has been developed.

Important parameters for audio conferencing and related applications are audio quality (audio bandwidth, distortion, noise), latency, bit rate and storage capacity. Latency requirements are mostly dependent upon whether the application is two-way (conference-like) or not (streaming-like). For conferencing, latency requirements may be as strict as 20 – 40 ms, as discussed above. While for streaming-like situations, the latency requirements may range from being important (keeping up with an ongoing event) to

almost non-existent (e.g. watching an old movie). In the first case latencies of seconds, or even minutes, may be acceptable. In the latter case, latency is only important and user noticeable during set-up. If the streaming may be ordered in advance, latency as long as the advance notices (hours, days) may be tolerated.

Available bandwidth is steadily increasing, and so is available storage capacity. What we see is an evolution where already established areas of use see an increase in available bit rate and storage capacity. Examples of this are network connections for home computers going from the use of modems (9.6 kb/s, 28.8 kb/s) via ISDN (64 kb/s) to ADSL (640/256 kb/s, 3000/500 kb/s), and PC hard drives, with storage capacities rising from tens of megabytes to hundred of gigabytes. At the same time, new areas and technologies arise, like mobile phones, mobile terminals and wireless networking. Although these new technologies typically start out with relatively low capacities, they also quickly evolve towards higher capacities.

Latency, however, is more directly bounded by fundamental physical laws and limitations. Signals will not travel faster than light, and the distance along the surface of the earth forms a practical lower limit for the length of the path between two endpoints. So, the theoretical lower limit to latency when transmitting a signal "half way round the world" is 67 ms ($20,000 \text{ km} / 3 \cdot 10^8 \text{ m/s}$). In practice, the latency is higher. The typical round-trip time (measured with "ping") between Trondheim and New Zealand is 320 ms, giving a one-way latency of 160 ms.

So, while available bandwidth and storage capacity is increasing, seemingly without limit, latency is bound by a limit to which we are actually quite close. And achievable latencies are, for many cases, as high, or higher, than applications requirements call for. In light of this, latency is likely to become the limiting factor for some classes of network communication technologies. This makes latency a very challenging aspect to work with.

The development of the AAC audio codec may serve as an example of improvement due to latency requirements. The original AAC (Advanced Audio Codec) has an algorithmic delay of several hundred milliseconds. The later AAC-LD (Low Delay AAC) has an algorithmic delay of 20ms, and a reported end-to-end implementation latency of about 45 ms [Hilpert00]. The later Ultra Low Delay codec from Fraunhofer sports an algorithmic delay of 6 ms [Hirsch04].

Quality metrics and perception

QoS may be defined as user satisfaction with the service, that is, the perceived quality by the end user for a specific service. This is a far-reaching concept since it involves direct effects such as perceived quality of the sound and video when multimedia content is streamed. Furthermore, indirect effects that affect the perceived quality could include price and expectation of the service, as well as the preconditioning of the customer. It is virtually impossible to derive a model that would predict a single-number perceived quality taking all these factors into account, in particular since many of the factors have strong individual weights. What is possible, however, is to quantify quality within well-defined subsets of this complex and with many of the factors controlled.

Subjective assessment of sound quality is used in two distinctly different situations: listening-only (streaming) and conversational (conference) situations. Even if the real application is a conversational situation, such as in telephony, the listening-only situation is often used in tests since such a test situation is easier to control. The result from a listening-only test does not translate perfectly to a conversational situation but a strong correlation exists.

For the listening-only situation, the most used subjective assessment method is a five-point category judgment where a scale of categories is used. For speech tests, this is the technique standardized by ITU [ITU-T P.800] and the quantity that results from such a test is called "mean opinion score", abbreviated MOS. Versions of this might compare a stimulus to a reference for an increased sensitivity to small differences, so-called Degradation Category Rating. Also for audio codec evaluation, the MOS test is standardized by ITU [ITU-R, BS.1116].

The MOS test has become a standard way to evaluate speech codecs and audio codecs. Using subjects is however time-consuming, so methods have been developed for objective evaluation of the sound quality of systems. One approach is to compare a degraded signal with a reference and to evaluate these two using some perceptual model. This so-called Full Reference (FR) approach is used in the PESQ method (Psychoacoustic Evaluation of Speech Quality), [ITU-T, P.862] and PEAQ (Perceptual Evaluation of Audio Quality), [ITU-R, BS.1387]. In these methods, a perceptual model of the hearing is used to determine the difference on hearing-related scales (for frequency and amplitude) between a reference signal and a signal that has been degraded by a system. A number of parameters have been adjusted to make the quality predictions fit as well as possible to a large data set of perception results from experiments with

human subjects. For the PESQ model it has been possible to reach a high correlation between the objective MOS value and MOS values from subjective evaluations and consequently it can be used for automatic evaluation of speech codecs. However, such an approach will always be less accurate for evaluating cases that are very different from the data set that was used for adjusting the model.

For communication over packet networks the effect of delays and packet losses is central. The so-called E-model includes the delay in an attempt to model the conversational quality [ITU-T, G.107] whereas PESQ does not. Studies of conversational quality have indicated quite variable results because the conversational style and the experimental setting will have much influence on the perceived degradation due to delays. The ITU recommendation is that a one-way delay should be below 150 ms for “essentially transparent interactivity” [ITU-T, G.114]. Still, some applications might be affected by even lower delays. As an example, for distributed music playing, which was referred to above, it is possible to measure an effect of musicians’ timing for delays as low as 20 – 40 ms.

In today’s world of multimedia communication over lossy networks like best-effort IP networks, it is crucial to be able to monitor the effects of compression- and transmission related distortions in order to quantify the user’s Quality of Experience (QoE) [Eberle 2005]. QoE relates the actual quality as perceived by an end-user to the overall communication system’s Quality of Service (QoS). Due to the nature of streaming media, quality monitoring has to be conducted in real time, and a reference stream for quality comparison is most often not available. In laboratory evaluations, however, the FR methods might be preferred. Thus one requires a No-Reference (NR) metric with the ability to estimate the end-user’s experience of a multimedia presentation without using an original audiovisual media stream as reference.

For IP networks, the deterioration in perceived quality is typically due to packet loss [Feamster 2002, Boyce 1998, Hillestad 2004, Bopardikar 2005]. The other major source of distortion and degradation of perceptual quality in multimedia communication is because of the inevitable coding and compression of media sources. In particular, for block-based video compression schemes such as the ISO/IEC and ITU standards (e.g. MPEG-1/2/4, H-261/3/4) the main forms of distortions include block impairment effects, blurring, ringing and the DCT basis image effect [Wu 1996, Yuen 1998].

NR metrics that have been proposed, in general try to quantify the effects of these distortions [Caviedes 2003, Marziliano 2002, Babu 2004] but the emphasis of research on NR metrics has been predominantly on quantifying the effects of block impairment artifacts [Wu 1998, Winkler 2001, Wang 2002, Gao 2002]. This is because block impairment artifacts tend to be perceptually the most significant of all coding artifacts. With the Video Quality Experts Group working towards their standardization [VQEG], NR metrics remain a topic of great research interest.

Challenges

The effects of delay and packet loss in packet-based transmission has introduced the challenge of handling delays and packet losses. Existing standards for speech coding and video conferencing have been extended to include some form of packet loss concealment but more development can be expected in this area. The delay that is introduced by the packet-based transmission is irrelevant for most one-way transmission applications but is crucial for two-way communication. From one end to the other end, the total delay includes the physical transmission delay as well as the buffer delays at both ends. These buffers are necessary because of delay variations, packet loss concealment algorithms, and coding/decoding processing delays. Notably, the average physical transmission delay might be a very small part of the total delay. Early demonstrations of distributed music playing across the North American continent used uncompressed audio and video to minimize the total delays since this application is very sensitive to delays.

For current multimedia communications there already exists a digital media market place where we all consume, be it our mobile phone, Internet based services or broadcasting. The major challenges arising are in transcoding between formats or event transmoding between media modalities such as speech, audio and video. This requires new ways of content representation, where the methods and functionalities must consider interactivity, and adaptive representations of the media. For quantifying quality we need quantitative measures for perception of digital media and the ability to measure this in IP networks and use the feedback from measurements to improve the quality in a dynamic way.

These challenges are depicted in Figure 3, where the enhanced version of the presented media makes use of available metadata for the media and its environment as well as quality metrics and other control signals given as feedback in the system.

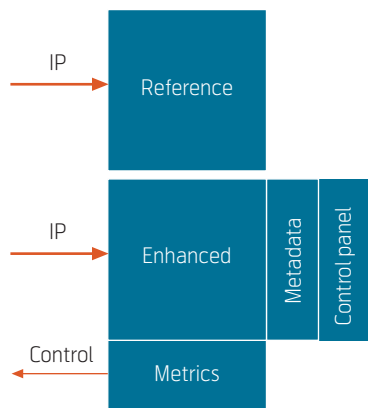


Figure 3 A rich controllable media viewer

Reference list (Labels in text refer to first authors' surname and year)

- Albanese, A, Blömer, J, Edmonds, J, Luby, M, Sudan, M. Priority Encoding Transmission. *IEEE Transactions of Information Theory*, 42 (6), 1996.
- Barbedo, J G A, Lopes, A. A New Cognitive Model for Objective Assessment of Audio Quality. *J. Aud. Eng. Soc.*, 53, 22, 2005.
- Bopardikar, A S, Hillestad, O I, Perkis, A. Temporal concealment of packet-loss related distortions in video based on structural alignment. In: *Proc. Eurescom Summit 2005*, Heidelberg, Germany, April 2005.
- Caviedes, J, Gurbuz, S. No-reference sharpness metric based on local edge kurtosis. In: *Proceedings of the International Conference on Image Processing*, Rochester, NY, September 22–25, 2002, 3, 53–56.
- Boyce, J, Galianello, R. Packet loss effects on MPEG video sent over public internet. In: *ACM International Multimedia Conference*, 1998, 181–190.
- Chen, P, Woods, J W. Bidirectional MC-EZBC with Lifting Implementation. *IEEE Transactions on Circuits and Systems for Video Technology*, 14 (10), 2004.
- Daniel, J, Nicol, R, Moreau, S. Further Investigation of High Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging. *Audio Eng. Soc. 114th Conv.*, 2003 March, Amsterdam, Preprint no. 5788.
- Eberle, W, Bougard, B, Pollin, S, Catthoor, F. From myth to methodology: Cross-layer design for energy-efficient wireless communication. In: *Proc. ACM/IEEE DAC*, Anaheim, USA, June 2005.
- Ebrahimi, T, Kunt, M. Visual Data Compression for Multimedia Applications. *Proc. IEEE*, 86 (6), June 1998.
- Feamster, N, Balakrishnan, H. Packet loss recovery for streaming video. In: *International Packet Video Workshop*, April 2002.
- Gao, W, Mermer, C, Kim, Y. A de-blocking algorithm and a blockiness metric for highly compressed images. *IEEE Transactions on Circuits and Systems for Video Technology*, 12 (12), 1150–1159, 2002.
- Hillestad, O I, Venkatesh Babu, R, Bopardikar, A S, Perkis, A. Video quality evaluations for UMA. In: *Proc. 5th International Workshop on Image Analysis for Multimedia Interactive Services*, Lisboa, Portugal, 21–23 April 2004.
- Hillestad, O I, Libæk, B, Perkis, A. Performance Evaluation of Multimedia Services over IP Networks. *IEEE Conference on Multimedia and Expo*, Amsterdam, The Netherlands, 6–8 July 2005.
- Hilpert, J, Gayer, M, Lutzky, M, Hirt, T, Geyersberger, S, Hoepfl, J. Real-Time Implementation of the MPEG-4 Low Delay Advanced Audio Coding Algorithm (AAC-LD) on Motorola DSP56300. *Audio Eng. Soc. 108th Conv.*, Paris, February 2000, Preprint 5081.
- Hirschfeld, J, Klier, J, Kraemer, U, Schuller, G, Wabnik, S. Ultra Low Delay Audio Coding with Constant Bit Rate. *Audio Eng. Soc. 117th Conv.*, San Francisco, October 2004, Preprint 6197.
- Johansen, S, Perkis, A. Unequal Error Protection for Embedded Codes over Parallel Packet Erasure Channels. Submitted to *IEEE Workshop on Multimedia Signal Processing*, Shanghai, China, October 2005.
- Kalman, M, Steinbach, E, Girod, B. Adaptive Media Payout for Low-Delay Video Streaming Over Error-Prone Channels. In: *IEEE Transactions on Circuits and Systems for Video Technology*, 14 (6), 2004.
- Marziliano, P, Dufaux, F, Winkler, S, Ebrahimi, T. A no-reference perceptual blur metric. In: *Proceedings of the International Conference on Image Processing*, Rochester, NY, 22–25 September 2002, 3, 57–60.
- ITU. *Methods for objective measurements of perceived audio quality*. Geneva, 2001. (ITU-R BS.1387-1)

- ITU. *Methods for subjective determination of transmission quality*. Geneva, 1996. (ITU-T P.800)
- ITU. *Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems*. Geneva, 1997. (ITU-R BS.1116-1)
- Mohr, A E, Riskin, E A, Ladner, R E. Graceful Degradation Over Packet Erasure Channels Through Forward Error Correction Codes. *Proceedings of Data Compression Conference*, IEEE, March 1999.
- ISO. *Information technology – Coding of moving pictures and associated audio for digital storage media up to about 1.5 Mbit/s – Part 2: Coding of moving pictures information*. 1991. (ISO/IEC JTC1 CD 11172, MPEG-1)
- ISO. *Information technology – Generic coding of moving pictures and associated audio information – Part 2: Video*. 1994. (ISO/IEC DIS 13818-2, MPEG-2)
- ISO. *Information technology – Coding of audio-visual object: Visual*. October 1997. (ISO/IEC JTC1 CD 14496-2 (MPEG-4))
- Ohm, J-R, van der Schaar, M, Woods, J W. Inter-frame wavelet coding – motion picture representation for universal scalability. In: *Signal Processing: Image communication*, 19 (2004), Elsevier, 877–908.
- ITU. *One-way transmission time*. Geneva, 2003. (ITU-T: G.114)
- ITU. *Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs*. Geneva, 2001. (ITU-T P.862)
- Puri, R, Ramchandran, K. Multiple Description Source Coding using Forward Error Correction Codes. *Proceedings of 33rd Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA. IEEE, 1999.
- Radha, H, van der Schaar, M, Chen, Y. The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming over IP. *IEEE Transactions on Multimedia*, 3 (1), 2001.
- RTP: *A Transport Protocol for Real-Time Applications*. January 1996. (RFC 1889)
- Stankovic, V, Hamzaoui, R, Xiong, Z. Packet Loss Protection of Embedded Data with Fast Local Search. *Proceedings of IEEE International Conference on Image Processing*, 2002.
- Stockhammer, T, Hannuksela, M M, Wiegand, T. H.264/AVC in Wireless Environments. *IEEE Transactions On Circuits And Systems For Video Technology*, 13 (7), 6657–6673, 2003.
- Sullivan, G J, Topiwala, P, Luthra, A. *The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions*. Presented at the SPIE Conference on Applications of Digital Image Processing XXVII Special Session on Advances in the New Emerging Standard: H.264/AVC, August, 2004.
- Sæbø, A. *Influence of reflections on crosstalk cancelled playback of binaural sound*. NTNU, Trondheim, 2001. (PhD thesis)
- Tamhankar, A, Rao, K R. An Overview of H.264 I MPEG4 Part 10. *EC-VIP-MC 2003, 4th EURASIP Conference focused on Video I Image Processing and Multimedia Communications*, Zagreb, Croatia, 2–5 July 2003.
- ITU. *The E-model, a computational model for use in transmission planning*. Geneva, 2003. (ITU-T. G.107)
- Yuen, M, Wu, H R. A survey of hybrid MC/DPCM/DCT video coding distortions. *Signal Processing*, 4 (11), 317–320, 1997.
- Venkatesh, B R, Bopardikar, A S, Perkis, A, Hillestad, O I. No-Reference metrics for video streaming applications, accepted in *The 14th International Packet Video Workshop (PV2004)*, 13–14 December 2004 at University of California, Irvine, USA.
- Video Quality Experts Group (VQEG)*. URL: <http://www.vqeg.org>
- Wang, Y, Zhu, Q-F. Error Control and Concealment for Video Communication: A Review. *Proceedings of the IEEE*, 86 (5), 1998.
- Wang, Z, Sheikh, H R, Bovik, A C. Noreference perceptual quality assessment of JPEG compressed images. In: *Proc. ICIP'02*, September 2002, 1, 477–480.
- Wiegand, T, Sullivan, G J, Bjøntegaard, G, Luthra, A. Overview of the H.264/AVC Video Coding Standard. *IEEE Transactions On Circuits And Systems For Video Technology*, 13 (7), 560–576, 2003.

Wenger, S. H.264/AVC Over IP. *IEEE Transactions On Circuits And Systems For Video Technology*, 13 (7), 645–656, 2003.

Winkler, S, Sharma, A, McNally, D. Perceptual video quality and blockiness metrics for multimedia streaming applications. In: *Proc. 4th International Symposium on Wireless Personal Multimedia Communications*, Aalborg, Denmark, September 2001, 553–556.

Woszczyk, W, Cooperstock, J, Roston, J, Martens, W. Shake, Rattle and Roll: Getting Immersed in Multisensory, Interactive Music via Broadband Networks. *J.Audio Eng. Soc.*, 53, 336–344, 2005.

Wu, H R, Hou, Y T, Zhu, W, Zhang, Y-Q, Peha, J M. Streaming Video over the Internet: Approaches and Directions. In: *IEEE Transactions on Circuits and Systems for Video Technology*, 11 (3), 2001.

Wu, H R, Yuen, M, Qiu, B. Video coding distortion classification and quantitative impairment metrics. In: *International Conference on Signal Processing*, 2, 962–965, 1996.

Wu, H R, Yuen, M. A generalized block-edge impairment metric for video coding. *IEEE Signal Processing Letters*, 70 (3), 247–278, 1998.

Other reading

Blauert, J. *Spatial Hearing, Revised ed.* Cambridge, MA, USA, The MIT Press, 1997.

Gibson, J D, Berger, T, Lookabaugh, T, Baker, R, Lindbergh, D. *Digital Compression for Multimedia*. Morgan Kaufmann, 1998.

Gilkey, R H, Anderson, T R. *Binaural and Spatial Hearing in Real and Virtual Environments*. Lawrence Erlbaum Associates, 1997.

Sun, M T, Reibman, A. *Compressed Video over Networks*. Marcel Dekker, 2000.

Special Issue Part One: Multimedia Signal Processing. *Proceedings of the IEEE*, 86 (5), 1998.

Special Issue Part Two: Multimedia Signal Processing. *Proceedings of the IEEE*, 86 (6), 1998.

Andrew Perkis was born in Norway 1961. He received his Siv.Ing. and Dr.Techn. degrees in 1985 and 1994, respectively. Since 1993 he has held the position of Associate Professor at the Department of Telecommunications at NTNU and as full professor since 2003. In 1999/2000 he was a visiting professor at The University of Wollongong, Australia. He is responsible for "Multimedia over IP networks" within the National Centre of Excellence – Quantifiable Quality of Service in communication systems at NTNU. Currently he is focusing on Multimedia Signal Processing and its use within The Multimedia Framework (MPEG-21), specifically; Creating advanced interactive media resources for multimedia communications, exploiting characterization and feed back from the IP network in codec design, perceptual metrics for measuring Quality of Experience and adaptive coding techniques.

email: andrew@q2s.ntnu.no

Peter Svensson was born in Sweden in 1964. He received his MSc and PhD degrees in 1987 and 1994, respectively, both from Chalmers University of Technology in Gothenburg, Sweden. Parts of a post doc position were spent at the University of Waterloo, Canada, and Kobe University, Japan. He has been professor in electroacoustics at NTNU since 1999. He is responsible for the research area "Audio over IP networks" within the National Centre of Excellence – Quantifiable Quality of Service in communication systems at NTNU, and project leader for "Acoustic Research Centre", a project funded by the Research Council of Norway. His research interests focus on sound reproduction and acoustic modelling in virtual environments (auralization), as well as perceptual aspects of reproduced audio and interaction over networks.

email: svensson@q2s.ntnu.no

Odd Inge Hillestad was born in Hønefoss, Norway in 1978. He received his Siv.Ing. degree from the Norwegian University of Science and Technology (NTNU) in 2002, where he has been working toward a PhD since January 2003. His research is being conducted at the Centre for Quantifiable Quality of Service in Communications Systems, with professor Andrew Perkis as the main supervisor. Research interests include video compression for packet networks and streaming media quality.

email: hillesta@q2s.ntnu.no

Stian Johansen was born in Harstad, Norway in 1978. He received his Siv.Ing. degree from the Norwegian University of Science and Technology (NTNU) in 2003, where he has been pursuing his PhD since. His research is being conducted at the Centre for Quantifiable Quality of Service in Communications Systems, with professor Andrew Perkis as the main supervisor. Research interests include 3D video coding, joint source/channel coding for packet networks and adaptive video communications systems.

email: stianjo@q2s.ntnu.no

Jijun Zhang received the MSc. Degree from Chalmers University of Technology, Gothenburg, Sweden in 2000. He is currently pursuing the PhD degree in multimedia signal processing in the Department of Electronics and Telecommunications, Norwegian University of Science and Technology, Trondheim, Norway. He was a research and teaching assistant in the Department of Applied Physics and Electronics, Umeå University in 2000, and a visiting researcher at the DISCOVER Lab, University of Ottawa, from September 2003 to September 2004. His research interests include multimedia adaptation, media conversion (transcoding) as well as video coding and image processing. Mr. Zhang is a student member of IEEE.

email: jijun@q2s.ntnu.no

Asbjørn Sæbø was born in 1968. He received his Siv.Ing. and Dr.Ing. degrees in 1995 and 2002. His general interests are research and development in audio, working within the areas of acoustics, electronics, music and IT. His doctoral work was on the reproduction of binaural sound. He has also worked with active noise cancellation (in the company Silence International) and as an independent consultant in audio and acoustics. He is currently employed at the Centre for Quantifiable Quality in Communication Systems, where he is working on distributed multimedia interaction within the audio over IP research area.

email: asbjorn.sabo@q2s.ntnu.no

Ola Jetlund was born in Norway 1974. He received his Siv.Ing. and Dr.Ing. degrees from the Norwegian University of Science and Technology (NTNU) in 1999 and 2005, respectively. Since 2004 he has held the position of University Lecturer at the Department of Telecommunications at NTNU. Currently he is focusing on Multimedia Signal Processing, wireless communications, feed back in IP networks to adjust scalable video codecs, and adaptive coding techniques to enhance Quality of Service in general.

email: ola.jetlund@q2s.ntnu.no

Peer-to-Peer IP Telephony

GEIR EGELAND AND PAAL ENGELSTAD



Paal E. Engelstad
is Research
Scientist at
Telenor R&D



Geir Egeland
is Research
Scientist at
Telenor R&D

The recent marriage between IP telephony and Peer-to-Peer communication, e.g. in applications such as Skype, has made a tremendous impact on the telecommunications industry. Although the original Internet design was based on a peer-to-peer architecture, applications such as Skype must rely on an overlay Peer-to-Peer network in order to cope with the underlying client-server paradigm imposed by NATs and firewalls. Initially peer-to-peer communication and IP telephony evolved independently. This article takes a closer look at the exciting development seen in the intersection between IP telephony and Peer-to-Peer communication. NATs and firewalls are identified as show-stoppers for further growth, while IPv6 might represent a solution.

Introduction

In August 2003 a small application called Skype was introduced on the Internet. It could be downloaded for free, and it enabled its users to make VoIP calls to each other free of charge. The most compelling feature of this application, in addition to being free, was that the voice quality was comparable to that of traditional fixed line telephony. Other applications such as MSN and Yahoo messenger had for some time offered its users voice communication, but with poor voice quality and with one major obstacle – they could not operate when the users were located behind a Network Address Translator. Skype on the other hand, introduced advanced traversal techniques enabling the users to communicate peer-to-peer even when both the caller and the callee are located behind a Network Address Translator.

After a general overview of peer-to-peer communication, we present IP telephony and how it can be deployed in peer-to-peer settings. The term “peer-to-peer” is not used as a euphemism for file sharing or other related activities, but in its original architectural sense, that all hosts on the network are logically equals. This was indeed the case in the early days of the Internet, but as the usage of the Internet changed, and the dominating communication form on the Internet became client/server, the end-to-end model of the Internet was broken, disabling peer-to-peer communication.

We also describe some of the technology used to traverse Network Address Translators to enable peer-to-peer communication, as well as some of the problems associated with these solutions.

Peer-to-Peer communication

The original design of the ARPANET, which forms the basis of the Internet architecture, was inherently peer-to-peer. Certainly, Internet connections differed

in bandwidth, latency, and reliability, but apart from those physical properties, any host could communicate on equal terms with any other host on the network. Any Internet host could provide any service to any other and access any service provided by them. This is opposed to the client-server paradigm, where communication only occurs between a client and a server. A peer-to-peer paradigm is more general, since it allows for communication between not only a client and a server, but between any client or any peers.

The first peer-to-peer applications that appeared were chatting applications on Unix operating systems. It allowed users to send instant messages to each other, similar to what Yahoo, AOL and MSN offer today. Since they did not need a centralized server to do this, unlike today’s *Instant Messaging* (IM) applications, the applications were truly peer-to-peer. They simply used Unix user names and the computer’s fully qualified domain name, i.e. by typing, “talk john_doe@hostname.somedomain.com” on the command line.

When the technology matured, the first peer-to-peer applications also offering voice appeared. One of these was *Speak Freely*, an application that allowed two or more people to conduct a real-time voice conference over the Internet or any other TCP/IP network.

Emerging problems with middleboxes

Due to a shortage of IPv4 addresses, Internet Service Providers were looking for techniques that preserved their IPv4 address space. At the same time, home users wanted to enable multiple machines to share the Internet connection. This was solved with the introduction of *Network Address Translation* (NAT) [1] [2], which permits multiple computers to share a common internet connection with a single IP address or with a limited pool of IP addresses.

At the time NATs were introduced, the vast majority of Internet communication was client-server-based, such as web-browsing, e-mail and file transfers. NATs were therefore designed according to this paradigm, and were embraced by the market because the deployment of peer-to-peer applications that required a strict peer-to-peer architecture was not considered significant. A network element that implements NAT is normally referred to as a “NAT-box” or simply as a “NAT” (Network Address Translator). The NAT has an “inside” encompassing the clients and an “outside” where all servers are located. In line with the client-server paradigm, the NAT assumes that the clients inside the NAT initiate all traffic. When a client initiates a session to a server outside the NAT, the NAT replaces the client’s private IP address with a global IP address of the NAT. The reply from the server is therefore sent to this address. When the reply is received by the NAT, it replaces the global IP address of the NAT with the client’s private IP address and forwards the packet to the client. All subsequent communication between the client and the server is subject to this kind of address translation.

The NAT approach assumes that all peers inside the NAT operate as clients and communicate with servers outside the NAT. A peer that is located outside a NAT cannot contact a peer located inside the NAT (although the opposite is possible). If two peers are located behind different NATs, peer-to-peer communication is not possible. Another problem is that NATs operate transparently to the peers. If the two peers cannot communicate because either one or both peers are located behind a NAT, the failure will not be detected by the peer-to-peer application.

The wide and rapidly growing deployment of NATs had direct influence on the development of peer-to-peer communication. After April 1996, for example, the development on the Speak Freely peer-to-peer application (mentioned above) was discontinued. The reason was the increasing usage of NATs, which destroyed the peer-to-peer architecture that the Internet was built upon.

Nowadays, NATs are so widely deployed that communication on the Internet that is peer-to-peer (in its strict architectural sense) cannot be guaranteed to work.

Overlay networks and Peer-to-Peer communication

The term peer-to-peer communication was re-introduced by an application called *Napster*. This was not an application for peer-to-peer communication using text or voice, but to let users share mp3-files over the Internet. After that, the term Peer-to-Peer (P2P) was

associated with file sharing. Napster provided a server where users could upload an index of the mp3-files available at the computer from which they connected to the network. The Napster clients could ask the server “*Where is this file*”, and the server would answer with the IP-address and port of the client that had provided the information. The requesting client could now contact the other Napster client directly and download the mp3-file, giving the clients the impression of communicating peer-to-peer. The Napster server was an easy target for legal authorities, and was soon forced to close its service.

Other file sharing applications were soon to follow where the centralized server of Napster was removed, enabling peers to form a distributed Peer-to-Peer overlay network (“P2P network”). The fundamental principle behind the P2P networks is that each and every node has equal importance in the network that is formed. Rather than a large number of client machines contacting one or more central servers, nodes interact directly with each other. Each node that participates in the P2P network provides server-like functionality and may act as both server and client within the system. The Gnutella Protocol [3] is an example of a technology used to realize a P2P network (Figure 1). In a Gnutella network, every client is a server, and vice versa. These so-called Gnutella servents perform tasks normally associated with both clients and servers. They provide client-side interfaces through which users can issue queries and view search results, while at the same time they also accept queries from other servents, check for matches against their local data set, and respond with applicable results. Due to its distributed nature, a network of servents that implements the Gnutella protocol is highly fault-tolerant, as operation of the network will not be interrupted if a subset of servents goes offline.

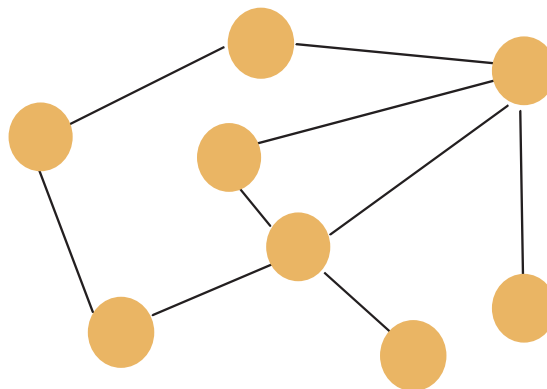


Figure 1 The architecture the Gnutella P2P overlay network

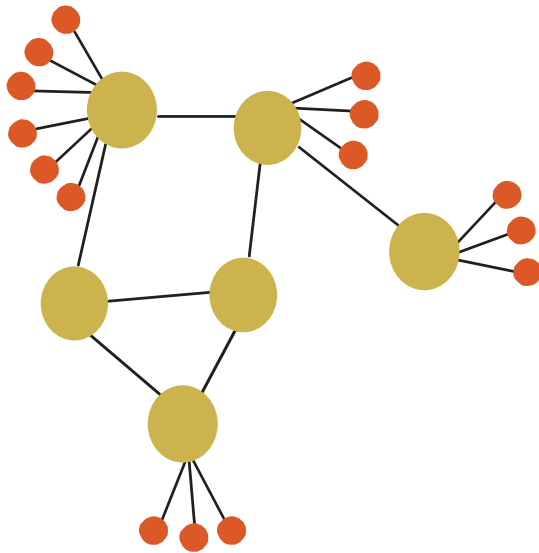


Figure 2 The architecture of the FastTrack P2P overlay network

With an overlay network, it is possible to implement distributed techniques to communicate freely, despite the possible presence of NATs in the network. This has made overlay networks an inevitable part of modern peer-to-peer communication in today's Internet. It should be noted that P2P networks are not peer-to-peer communication in the strict architectural sense. That is, underneath the P2P network peers that are located inside a NAT might always operate as clients, while peers not located inside a NAT might always operate as servers. However, for the P2P network, on the contrary, this underlying networking is concealed. In the overlay network, all nodes are peers that operate on equal terms. In this paper, we use the terms "P2P" or "Peer-to-Peer" (with capital "P"s) to distinguish overlay network architectures from network architectures that are "peer-to-peer" (with lower-case "p"s) in its architectural sense.

The most popular P2P networks today are overlay networks that use the second-generation P2P protocol FastTrack [4]. The FastTrack network was based on the Gnutella protocol [3], but was extended with the addition of *Supernodes* to improve scalability (Figure 2). Since most P2P-users have a higher download rate than upload rate, their links often became congested with search queries. To avoid this, P2P-users with such connections can upload its index of offered files to a supernode. The supernode can then answer search requests on behalf of the P2P-user with the slow connection. The supernodes in a P2P network can also act as proxy for P2P-users behind a NAT, since two P2P-users, each located behind their respective NAT, have no means of communicating peer-to-peer. The supernode functionality is built into the client software, and a P2P-user with a fast network connection and a computer with a global routable IP-address can become a supernode.

IP Telephony as a peer-to-peer application

Nowadays, end-to-end IP telephony comes in two forms. One form is where the end-user subscribes to a VoIP service from an IP telephony operator. The user uses some IP telephony equipment that is connected to the user's Internet connection, and the operator implements some IP telephony infrastructure on the Internet that the user connects to. This resembles traditional telephony in the way that the subscriber connects to a predefined network infrastructure. It also resembles the client-server paradigm with the users acting as clients and the IP telephony infrastructure acting as the server.

Another form of IP telephony is Peer-to-Peer. The IP telephony infrastructure mainly consists of the overlay network formed by all the other IP telephony peers. This is the main focus of this article. In the following we will present a number of VoIP Peer-to-Peer applications.

Skype

The Peer-to-Peer application Skype [5] is an example of a VoIP application that combines the Peer-to-Peer ability of modern P2P-networks with telephony. Skype operates over the FastTrack P2P-network, and takes advantage of the supernode functionality that effectively provides relay servers for Skype users behind a NAT. Skype users can also call traditional telephone numbers for a fee (SkypeOut) or receive calls from traditional phones (SkypeIn). SkypeOut and SkypeIn are realized by implementing gateways between the Skype Peer-to-Peer network and traditional PSTN telephony networks.

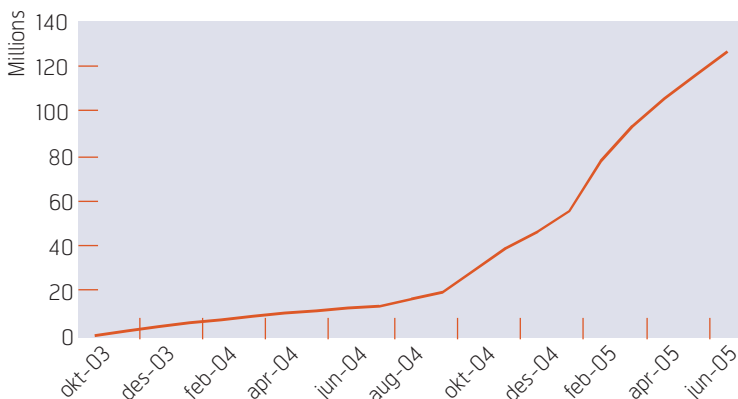


Figure 3 Downloading rate of Skype

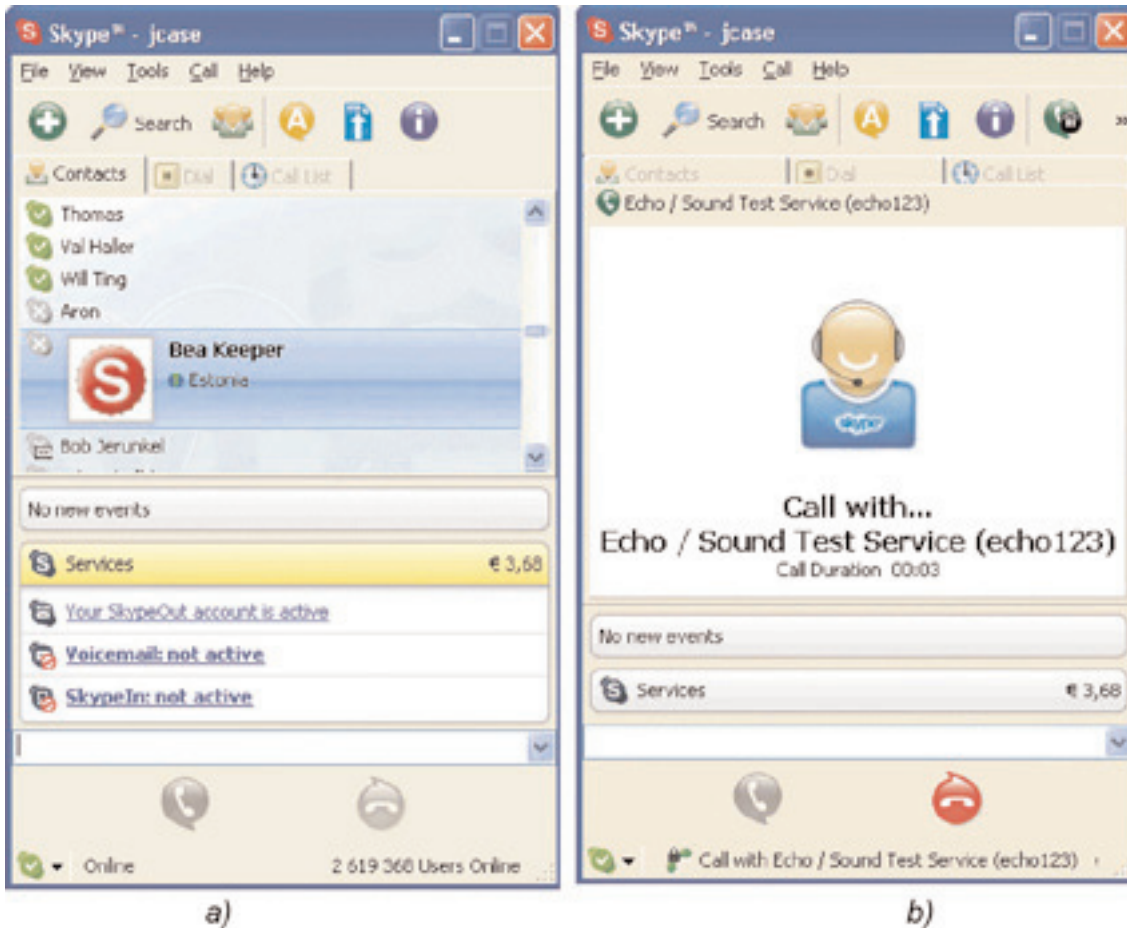


Figure 4 The Skype user interface. a) Contacts and Event, b) Talking

Skype is the fastest growing application in the history of the Internet. The downloading rate is shown in Figure 3. As other instant messaging (IM) applications, Skype has capabilities for voice-calls, instant messaging, audio conferencing, and buddy lists.

The Skype protocol is proprietary and encrypted and is thus not public knowledge. The Skype application has been studied in [6] where an attempt to decrypt its protocol was made. It was discovered that there are three types of nodes in the Skype overlay net-

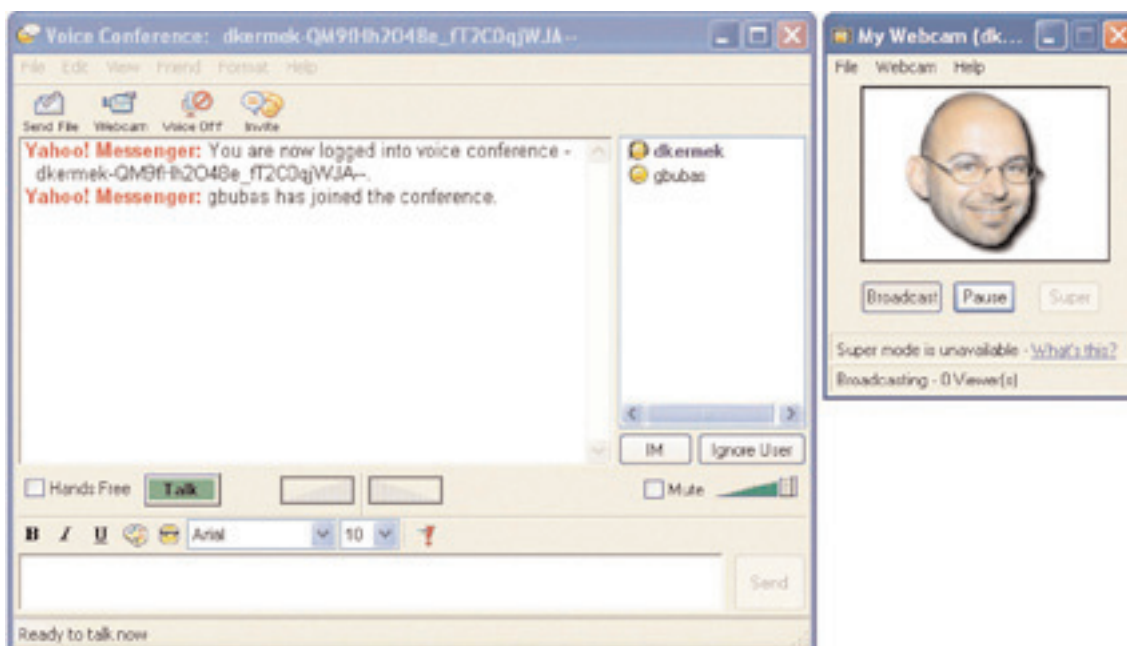


Figure 5 Yahoo Messenger user interface

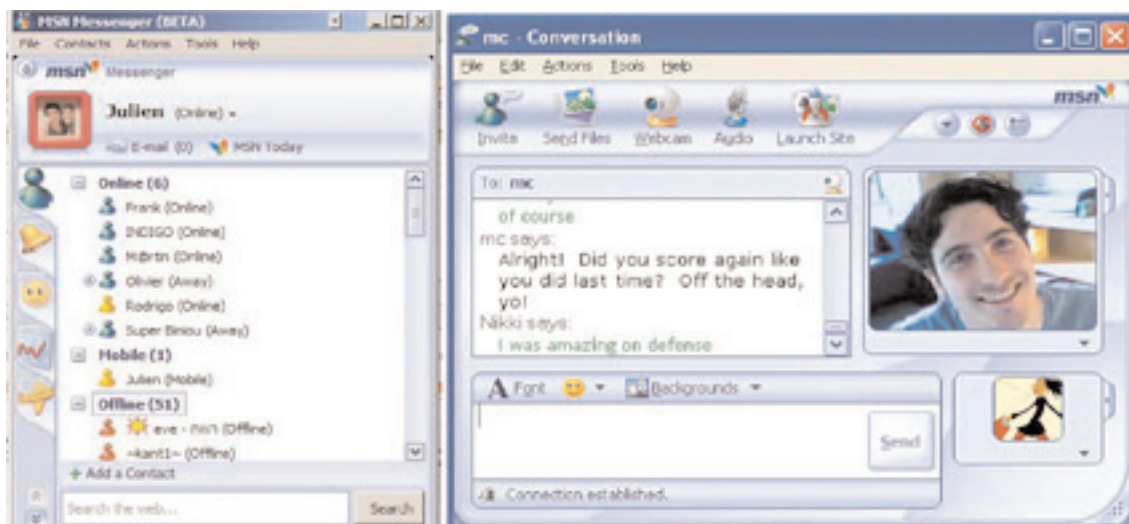


Figure 6 MSN Messenger user interface

work: ordinary hosts, supernodes and a login server where user names and passwords are authenticated and stored. An ordinary host must connect to a supernode to register with the Skype login server. All communications between any pair of Skype users consisting of any combination of voice, video, text chat or file transfer are carried over an encrypted Skype “session layer” that is established between the communicating users before messaging begins.

The Skype user interface is illustrated in Figure 4.

Yahoo Messenger

Yahoo Messenger is an instant messenger application provided by Yahoo and has been widely available for many years. It is heavily integrated with Yahoo’s other services, such as online shopping, web search, multi-user online games and email. In its latest release it has included a VoIP service. In earlier versions Yahoo Messenger also offered voice communication, but only between computers where only one of the users was located behind a NAT. Yahoo Messenger now offers VoIP also when both users are located behind NATs.

In June 2005 Yahoo acquired a VoIP company called DialPad. This company has technology for connecting VoIP calls on the Internet with the PSTN network. It is assumed that this technology will be integrated with Yahoo Messenger. The Yahoo Messenger user interface is illustrated in Figure 5.

MSN Messenger

MSN Messenger is Microsoft’s instant messaging client for Windows computers. The major use of the software is for instant messaging, although other features which now come as standard include support for

voice conversations, full screen audio video conversations, transferring files, and built-in multi-user online games.

The protocol used by MSN Messenger is closed, but it is common knowledge that the basic elements are as follows [7]:

Notification Server (NS): The connection to a notification server is the basis of an MSN Messenger session, as it handles the users’ presence information. If you are disconnected from the notification server, you are no longer online to your buddies. The main purpose of the notification server is to handle presence information about the user and the principals whose presence he has subscribed to.

Switchboard (SB): The switchboard handles instant messaging sessions between principals. In other words, each person in an MSN chat corresponds to a connection to a shared switchboard session. Invitations to other services such as file transfer and Net-Meeting are also sent and received through the SB.

In August 2005 Microsoft bought a small VoIP company called Teleo [8]. This company has a VoIP application that is able to connect to the public telephony network. Most likely this technology will be integrated with MSN Messenger, providing the same VoIP capabilities as Skype. Even though Skype has had a huge success, it is nowhere near MSN Messenger when it comes to the number of installations.

The MSN Messenger user interface is illustrated in Figure 6.

Google Talk

The newest contender on the IM/VoIP arena is Google. In August 2005 Google announced Google Talk, which is a small IM application that is integrated with Google Mail and offers VoIP communication between other Google Talk users. The Google Talk user interface is illustrated in Figure 7.

SIP in a peer-to-peer telephony setting

In order to realize end-to-end IP telephony, in terms of initiating and terminating calls, the IETF developed the *Session Initiation Protocol* (SIP) [9]. It allows for direct signalling between two end-users, where one user can call another by sending an “invite” message using a URI on the form `SIP:bill@online.no` (Figure 8). The message contains the IP address, protocol (typically RTP/UDP), port number and possible codecs preferred by the caller. The callee sends a reply, containing its own address, port number and the codec. When this negotiation is complete, the reply is finally acknowledged (Figure 8), and an RTP session is set up in each direction to carry the actual telephone conversation.

Although SIP is designed to work in a strict peer-to-peer setting (as shown in Figure 3), it is normally used in combination with a SIP server. This is shown in Figure 9. (Here, only the initial invite message and the resulting media session are shown.)

SIP fits well in a peer-to-peer setting, since it is designed for direct communication between the clients. First, all servers in SIP are optional. Second, even when a server such as a proxy server is utilized, the packets are routed on a peer-to-peer basis after the initial exchange. Although SIP is based on an inherent peer-to-peer architecture, most applications using SIP today register their IP addresses with the SIP server so that the other users can reach them. In fact, the use of SIP servers may also be used to realize certain forms of mobility.

Although SIP was originally designed with IP telephony in mind, it might in principle be used to set up all kinds of communication sessions, such as chat sessions, video conferencing, and so forth. Furthermore, instance messaging and presence service (e.g. in terms of buddy-lists) are features that are now available with SIP. Instant messaging and presence service are commonly implemented by a large number of different peer-to-peer applications. Hence, SIP is a very suitable candidate technology both for IP telephony and peer-to-peer networking.

Peer-to-Peer systems inherently have high robustness and fault tolerance, because the network is self-organizing without the use of centralized servers. A sensi-

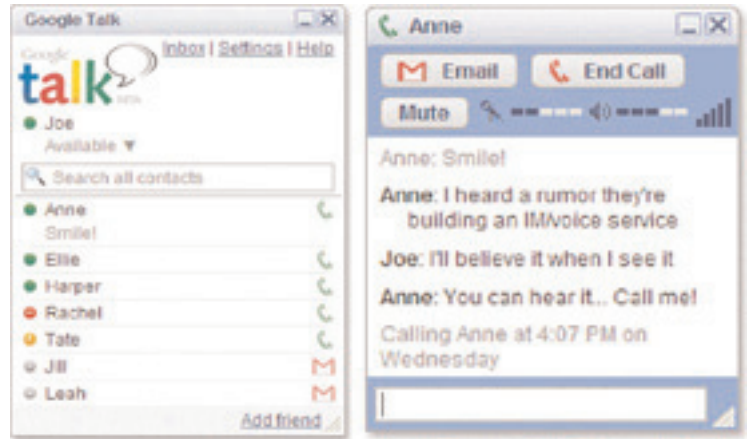


Figure 7 Google Talk user interface

ble approach might be a hybrid solution that will benefit from the scalability offered by centralized systems such as SIP and the reliability offered by P2P. A solution using P2P-technology in combination with SIP is suggested in [10][11]. Unlike a conventional SIP architecture, the P2P SIP system requires no central servers, and the peers connect directly to a few other peers, forming an overlay network. P2P enabled SIP nodes can then communicate with other P2P enabled SIP nodes to establish sessions.

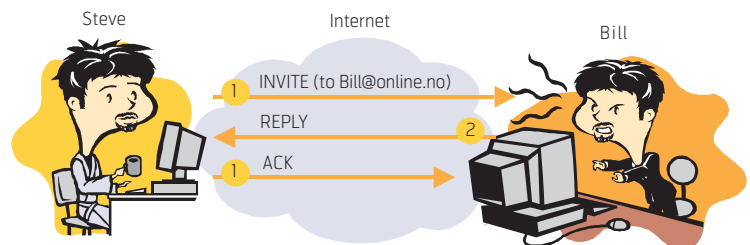


Figure 8 The SIP signaling (peer-to-peer)

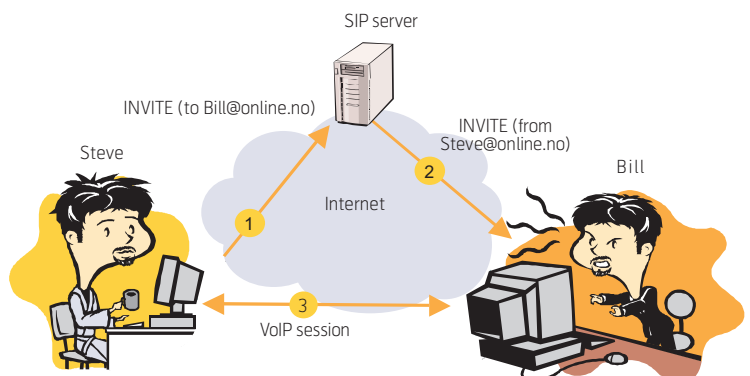


Figure 9 SIP signaling using a SIP server

A shortcoming of SIP is when both users are located behind a NAT, SIP cannot manage to establish the VoIP session without taking special measures to traverse the NAT. The problem of NAT traversal will be discussed in more detail later in the next chapter.

Enablers and showstoppers for peer-to-peer communication

Network Address Translation (NAT)

The P2P applications, SIP and others, all experience some problems when there is a NAT in the communication path. The operation of a NAT is to be an active unit placed in the data path, usually as a functional component of a border router or site gateway. A NAT intercepts all IP packets, and may forward the packets onward with or without alteration of the contents of the packets, or it may choose to discard the packets. The essential difference from a conventional router or a firewall is the ability of the NAT to alter an IP packet before forwarding it. NATs are similar to stateful firewalls, and different from routers, in that they are topologically sensitive. They have an “inside” and an “outside,” and undertake different operations on intercepted packets depending on whether the packet is going from inside to outside, or in the opposite direction.

The header of an IP packet contains the source and destination IP addresses. If the packet is being passed in the direction *from* the inside *to* the outside, a NAT rewrites the source address in the packet header to a different value, and alters the IP and TCP header checksums in the packet at the same time to reflect the change of the address field. When a packet is received *from* the outside destined *to* the inside, the

destination address is rewritten to a different value, and again the IP and TCP header checksums are recalculated. The “inside” does not use globally unique addresses to number every device within the network served by the NAT. The inside (or “local”) network may use addresses from private address blocks, implying that the uniqueness of the address holds only for the site. The operation of a NAT is illustrated in Figure 10.

A variant of the NAT is the *Port-Translating NAT*, or NATP. This form of NAT is used in the context of TCP and *User Datagram Protocol* (UDP) sessions, where the NAT maps the local source address and source port number to a public source address and a public-side port number for outgoing packets. Incoming packets addressed to this public address and port are translated to the corresponding local address and port.

There are many types of NATs, and different types of NATs require different solutions for NAT traversal. Generally, one can divide NATs into:

Symmetric NAT: Here the NAT mapping refers specifically to the connection between the local host address and port number and the destination address and port number and a binding of the local address and port to a public side address and port. This is the most restrictive form of NAT behavior under UDP, and it has been observed that this form of NAT behavior is becoming quite rare, because it prevents the operation of all forms of applications that undertake referral and handover.

Full-cone NAT: A full-cone NAT is the least restrictive form of NAT behavior. It implies that any remote

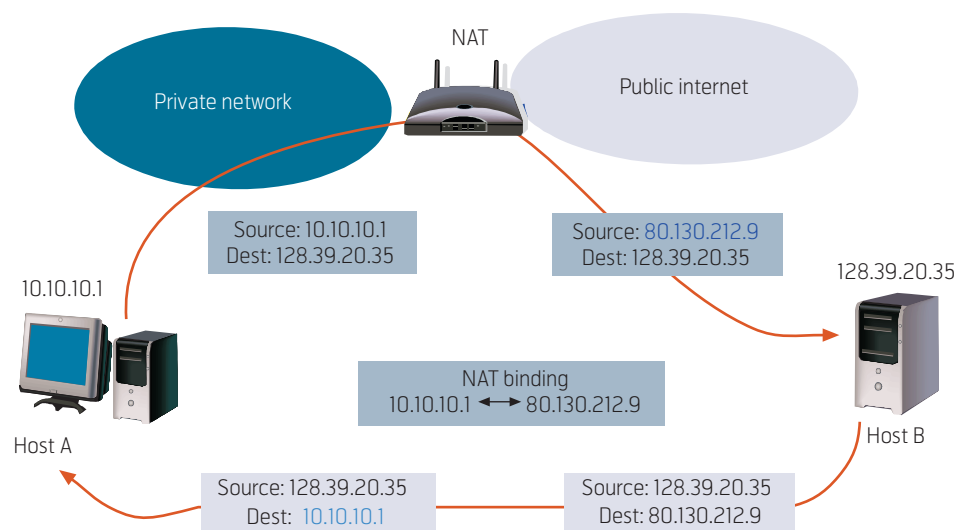


Figure 10 Operation of a NAT

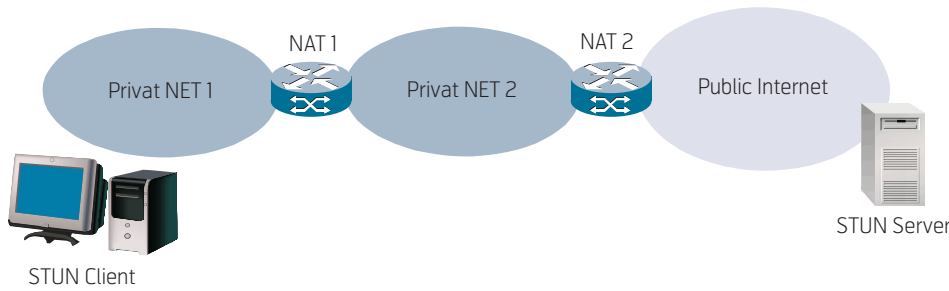


Figure 11 STUN Scenario

host on any remote port address can use the binding of a local address and port to a public-side address and port, when the binding has been established.

Restricted-cone NAT: A restricted-cone NAT is one where the NAT binding is accessible only by the destination host, although in this case the destination host can send packets from any port address after the binding is created.

Port-restricted-cone NAT: A port-restricted-cone NAT is one where the NAT binding is accessible by any remote host, although in this case the remote host must use the same source port address as the original port address that triggered the NAT binding.

Since the common use of IP was dominated by client/server communication, NATs were unfortunately designed under the assumption that no communication is peer-to-peer. These middleboxes assume that the server is located on the publicly routed Internet. However, as applications developed, it soon became clear that solutions for NAT traversal were necessary.

Simple Traversal of UDP through NATs (STUN)

For applications to be able to discover if they are behind a NAT, the *Simple Traversal of UDP through NATs* (STUN) protocol [12] was developed. STUN is a probe system that examines the interchange between a STUN client that may lie behind a NAT and a STUN server that is positioned on the public side of the NAT. The protocol is a simple UDP request-response protocol that uses embedded addresses in the data payload, and compares these addresses with header values in order to determine the type of NAT that may lie in the path between the client and server. The basic operation of STUN is to use a common request of the form: "Please tell me what public address and port values that were used to send this query to you." A typical STUN scenario is shown in Figure 11.

STUN attempts to determine the type of NAT by a structured sequence of requests and responses. This sequence is illustrated in Figure 12. The client sends an initial request to the STUN server. If the public address and port in the returned response are the same as the local address, then the client can conclude that there is no NAT in the path between the client and the server. If the values differ, the client can conclude that there is a NAT on the path. STUN then uses subsequent requests to determine the type of NAT.

Traversal Using Relay NAT (TURN)

STUN allows a client to obtain a transport address (and IP address and port) that may be useful for receiving packets from a peer. However, addresses obtained by STUN may not be usable by all peers. Those addresses work depending on the topological conditions of the network. Therefore, STUN by itself cannot provide a complete solution for NAT traversal.

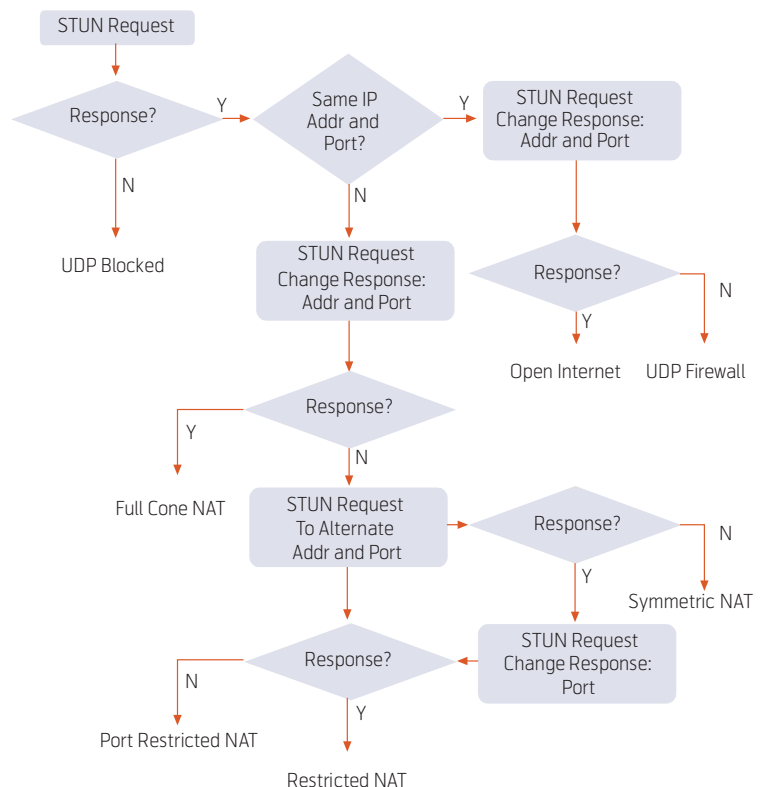


Figure 12 NAT discovery process using STUN

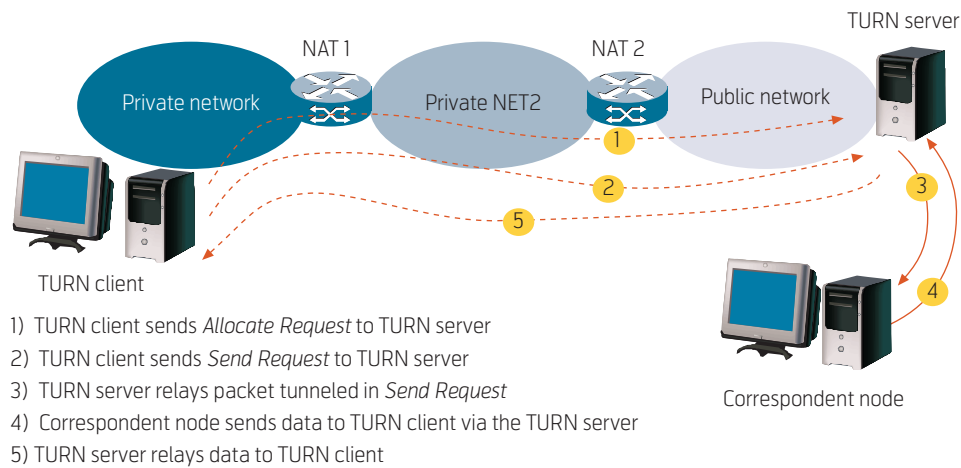


Figure 13 Operation of TURN

A complete solution requires that a client obtain a transport address that is commonly available from the public Internet. This can be accomplished by relaying incoming data to the client through a server that resides on the public Internet.

Traversal Using Relay NAT (TURN) [13] is a protocol that allows for an element behind a NAT or firewall to receive incoming data over TCP or UDP connections. It is most useful for elements behind symmetric NATs or firewalls that wish to be on the receiving end of a connection to a single peer. TURN does not allow for users to run servers on well-known ports if they are behind a NAT. It supports the connection of a user behind a NAT to only one single peer. In that regard, its role is to provide the same security functions as provided by symmetric NATs and firewalls, but to *turn* them into port-restricted NATs.

Unlike a STUN server, a TURN server provides resources to clients that connect to it. Therefore, only authorized clients can gain access to a TURN server. This requires that TURN requests are authenticated. TURN assumes the existence of a long-lived shared secret between the client and the TURN server in order to achieve this authentication. The client uses this long-lived shared secret to authenticate itself in a *Shared Secret Request*, sent over *Transport Layer Security (TLS)*. The *Shared Secret Response* provides the client with a one-time username and password. The TURN protocol is illustrated in Figure 13.

NAT Traversal of peer-to-peer IP telephony

Skype provides a solution for NAT and firewall traversal, and it is assumed that the Skype client uses a variation of the STUN and TURN protocols to determine the type of NATs and firewalls it is behind.

In [6] it was found that the Skype client is initialized with a set of bootstrap supernodes. These are stored in a host cache and contain IP-addresses / port pairs of active supernodes. A Skype client will learn the IP address of a login server from one of the supernodes, and will attempt to login there. It is not known whether it is the supernodes or the login server that act as a STUN/TURN server. It might even be some of the ordinary nodes themselves. It is shown in [6] that when the caller and the callee are both behind port restricted NATs, the traffic is relayed by another online Skype node.

It is well known that Yahoo and MSN Messenger are also using a STUN/TURN alike technology.

IPv6 as the future of peer-to-peer communication

NAT boxes are rapidly being deployed in combination with IEEE 802.11 wireless technology. As more and more peer nodes are located behind a NAT, the number of available peers with a public IP address will decrease. If the rate of new Skype users behind a NAT becomes higher than the rate of new Skype users with a public IP address, peer-to-peer telephony services such as Skype will be forced to deploy relay servers to maintain the same level of availability. This will transform the peer-to-peer architecture of Skype into a traditional client/server architecture. This new infrastructure will come with a price tag in the form of server boxes, bandwidth, maintenance and management. This is also the case for Yahoo and MSN Messenger.

Since the IPv4-based Internet is no longer peer-to-peer, there will be a need for middleboxes that bypass problems caused by NATs and firewalls. Some of the problems peer-to-peer services will face caused by this are:

- STUN and TURN servers are single point of failures in a network.
- STUN/TURN servers need maintenance and management
- STUN/TURN servers are vulnerable to DoS attacks
- STUN and TURN workarounds are per-application solutions. Since applications are not modular, every application will need to have its own solution for NAT and firewall traversal.

One commonly held belief is that the deployment of IPv6 [12] will eliminate the problem of NATs within the Internet. There will be plentiful public IPv6 address space available and no particular reason to deploy NATs in an IPv6 realm. That does not say that IPv6 NATs will not be implemented, nor used. Indeed IPv6 NATs are already available, and they are being used, albeit to a small extent. NATs are, rightly or wrongly, considered to be part of a security solution for a site because of their filtering properties that prevent incoming packets from entering the site. So it is reasonable to predict that some use of NAT will be seen in IPv6, although the level of IPv6 NAT will probably not be anywhere near that of NAT in IPv4.

IPv6 provides the necessary address space so that NATs are unnecessary. Hence, by converting to IPv6, the client/server communication paradigm that has been forced upon us by the deployment of NATs will evaporate. One of the most compelling consequences of a transition to IPv6 is that the simple and powerful Internet model is restored. Not only will IPv6 allow for the end-to-end principle in Internet communication. First and foremost, it will restore the peer-to-peer communication paradigm on the Internet. This will pave the way for a series of peer-to-peer services, which will for certain include voice communication. When IPv6 is deployed, Skype, MSN and Yahoo can easily be upgraded to support IPv6. IPv6 will be able to reduce the complexity of these VoIP applications, since there will no longer be any need for STUN/TURN solutions. The need for relaying of VoIP sessions will also be reduced to a minimum, since all users will be able to communicate peer-to-peer.

IPv6 will also reduce the complexity of developing peer-to-peer applications and services. Hopefully this will stimulate the development of other peer-to-peer voice services. Microsoft has stated that they see IPv6 as an enabler for peer-to-peer services [15] and are already supporting IPv6 in all their latest operating systems. They also provide APIs and development kits for P2P applications/services [15].

Concluding remarks

There is no doubt that 'Voice over IP' telephony is growing rapidly, and it is equally certain that it will grow even faster as the solutions continue to develop and as ever-expanding populations of potential users come into contact with the concept. Skype has pretty much dominated the VoIP scene the last year or so, as IP telephony finally hit the Internet. However, Skype has the following properties that may cause problems in the future:

- The protocol is proprietary, unlike open standards such as SIP.
- It provides a single service that makes calls or instant messages and does not provide an architecture for new services.
- It depends on the availability of public IPv4 addresses

Skype's VoIP technology is based on software from the company Global IP Sound (GIPS) [16]. This software is also available to others. Other VoIP applications supporting this technology are Google Talk [17] and MSN Messenger [18]. Gizmo [19] is another free VoIP application based on GIPS, and while Skype is proprietary, Gizmo has been built using an open source philosophy around the emerging SIP standards. In addition to being based on the SIP open standard, Gizmo is also committed to interconnecting its IP telephony system with those operated by other organizations. As more VoIP applications become available, Skype will face serious challenges, especially as SIP matures.

All peer-to-peer VoIP services will, however, have to deal with NATs. Despite their shortcomings and despite the problems NATs create for numerous applications, many NATs are deployed in the IPv4 realm. However, it is highly unlikely that a NAT-based architecture will scale to support peer-to-peer services, and IPv6 is therefore a prerequisite for a continued growth of peer-to-peer communication.

References

- 1 Egevang, K, Francis, P. *The IP Network Address Translator (NAT)*. May 1994. (RFC 1631)
- 2 Srisuresh, P, Egevang, K. *Traditional IP Network Address Translator (Traditional NAT)*. January 2001. (RFC 3022)
- 3 *The Gnutella Protocol*. March 10, 2006 [online] – URL: http://www9.limewire.com/developer/gnutella_protocol_0.4.pdf

- 4 *FastTrack Protocol*. March 10, 2006 [online] – URL: <http://en.wikipedia.org/wiki/FastTrack>
- 5 *The Skype VoIP application*. March 10, 2006 [online] – URL: www.skype.com
- 6 Baset, S A, Schulzrinne, H. *An Analysis of the Skype Peer-to-peer Internet Telephony Protocol*. March 10, 2006 [online] – URL: <http://arxiv.org/ftp/cs/papers/0412/0412017.pdf>
- 7 *MSN Messenger Protocol*. April 27, 2006 [online] – URL: <http://www.hypothetic.org/docs/msn/index.php>
- 8 *MSN Teleo*. March 10, 2006 [online] – URL: <http://teleo.msn.com/>
- 9 Rosenberg, J, Schulzrinne, H et al. *SIP: Session Initiation Protocol*. June 2002. (RFC3261)
- 10 Bryan, D, Jennings, C. *A P2P Approach to SIP Registration*. [online] – URL: <http://www.ietf.org/internet-drafts/draft-bryan-sipping-p2p-02.txt> (work in progress, published 5 March 2006)
- 11 Johnston, A. *SIP, P2P, and Internet Communications*. [online] – URL: <http://www.ietf.org/internet-drafts/draft-johnston-sipping-p2p-ipcom-02.txt> (work in progress, published March 5, 2006)
- 12 Rosenberg, J, Weinberger, J, Huitema, C, Mahy, R. *STUN – Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)*. March 2003. (RFC 3489)
- 13 Rosenberg, J, Mahy, R, Huitema, C. *Obtaining Relay Addresses from Simple Traversal of UDP Through NAT (STUN)*. [online] – URL: <http://www.ietf.org/internet-drafts/draft-ietf-behave-turn-00.txt> (published February 27, 2006)
- 14 Deering, S, Hinden, R. *Internet Protocol, Version 6 (IPv6) Specification*. December 1998. (RFC2460)
- 15 *Windows Peer-to-Peer Networking*. March 10, 2006 [online] – URL: <http://www.microsoft.com/windowsxp/p2p/default.msp>
- 16 *Global IP Sound*. March 10, 2006 [online] – URL: <http://www.globalip.com/>
- 17 *Voice codecs in Google Talk*. March 10, 2006 [online] – URL: <http://www.google.com/talk/developer.html#codecs>
- 18 *Microsoft and Global IP Sound Partner*. March 10, 2006 [online] – URL: <http://blog.tmcnet.com/blog/tom-keating/voip/microsoft-and-global-ip-sound-partner.asp>
- 19 *The Gizmo Project*. March 10, 2006 [online] – URL: <http://www.gizmoproject.com/>

Paal E. Engelstad completed his PhD on resource discovery in Mobile Ad hoc and Personal Area Networks in 2005. He has also a Bachelor and Masters degree (Honours with Distinction) in Applied Physics from NTNU, Norway, and a Bachelor degree in Computer Science from University of Oslo, Norway. After working five years in industry, he joined Telenor R&D where he focuses on IETF and IP technology (e.g. IP mobility, IPv6, QoS, MANET and AAA-issues) and IEEE wireless technologies (e.g. 802.11, 802.15 and 802.16). Paal Engelstad has published 28 refereed papers and holds three patents (two pending).

email: Paal.Engelstad@telenor.com

Geir Egeland holds a B.Eng (Hons) from the University of Bristol and has for the last ten years worked as a research scientist in the field of mobile networks. He is currently with Telenor R&D where his work is mainly focused on mobility for IP networks, with a particular emphasis on MANET and IPv6. Geir Egeland was formerly employed by the Norwegian Defence Research Establishment (NDRE) as a research scientist working on design and analysis of MAC and routing protocols for mobile ad hoc network.

email: geir.egeland@telenor.com

Voice over IP in the context of 3G mobile communications

INGE GRØNBÆK



Inge Grønbaek is Senior Adviser at Telenor R&D

The 3GPP architecture required for a global Voice over IP (VoIP) service is described. This shows that extensive interworking arrangements are required, resulting in an exceedingly complex architecture. Harmonization of the network architecture, with a common core, is therefore highly desirable and proposed as the bearer for a global SIP based service control as provided by the IP multimedia Sub-system (IMS). The offering of access independent services, by applying IMS for cellular and wireline networks, is a first step driven by ETSI TISPAN, providing a common converged fixed mobile architecture. Such an IMS overlay shields applications from the specifics of the different network and access technologies. This is the first step towards a harmonized architecture, but further gains can and need to be achieved by the agreement on functionality and protocols for the proposed common converged IP based core network. This asks for agreement on common solutions, e.g. for security, for global mobility management, for QoS control, for multicast, for codecs etc. The common core allows interworking to be carried out on the edge, avoiding the N-square interworking scenarios. Selecting a target core technology will implicitly define the direction of migration while still maintaining the freedom of innovation. The agreement on such a core will refocus the competition from being system oriented (i.e. competition between the cellular, fixed and broadcast families of systems) towards competition on services via new access technologies. The proposal will additionally give operators the opportunity to clean up and take control over the plethora of administrative operations support systems thereby reducing the operational and capital expenditure.

1 Introduction

This article describes aspects and challenges for Voice over IP (VoIP) implementation in the context of 3G (i.e. cellular infrastructures interworking with fixed networks including the Internet). The 3G architecture of most popularity, e.g. in Europe is defined by the third generation partnership project (3GPP). The 3GPP reference architecture is defined in [1], and a very brief presentation of its functional components is given in Appendix A.

The presentation is roughly structured as follows:

- Architecture for VoIP with key functions of the IP multimedia subsystem (IMS)
 - SIP for session control
 - Interworking services across multiple domains
- Next Generation Network (NGN) development
- Core architecture for harmonisation of technologies
- Concluding remarks

The first part introduces the 3GPP architecture and the functional components required for a global VoIP service. From a 3G point of view a separate plain VoIP service is of limited or no interest since the existing cellular infrastructures already are specialised for voice applications. The rationale for introducing VoIP in 3G systems is to allow introduction of additional service features, and to allow integration of voice as a component of new multimedia services.

A key topic for this overview is therefore the IP multimedia subsystem. IMS provides mechanisms for control of multimedia services. IMS may be used as an overlay both to cellular and fixed networks, and offers the capability to act as a vehicle for fixed mobile convergence (FMC).

An important part of the document covers the required interworking. This is essential to obtain ubiquity in VoIP services. SIP, H.323 and circuit domain interworking cases are included.

The next part of the document gives some notes regarding VoIP NGN related activities of Standards Developing Organizations (SDO). This includes combinational services for adding one or more IP multimedia component(s) to a Circuit Switched (CS) call, and Packet Switched (PS) to CS domain handover. Such services are required since certain 3G networks do not provide the required packet domain bearer quality for VoIP.

The last part introduces the notion of a target core architecture that will simplify interworking, network development and provide direction for innovation.

2 IMS subsystem

The 3GPP, ETSI and Parlay Forum have jointly defined IMS [2]. Originally IMS was motivated from a vision of widespread support of multimedia services

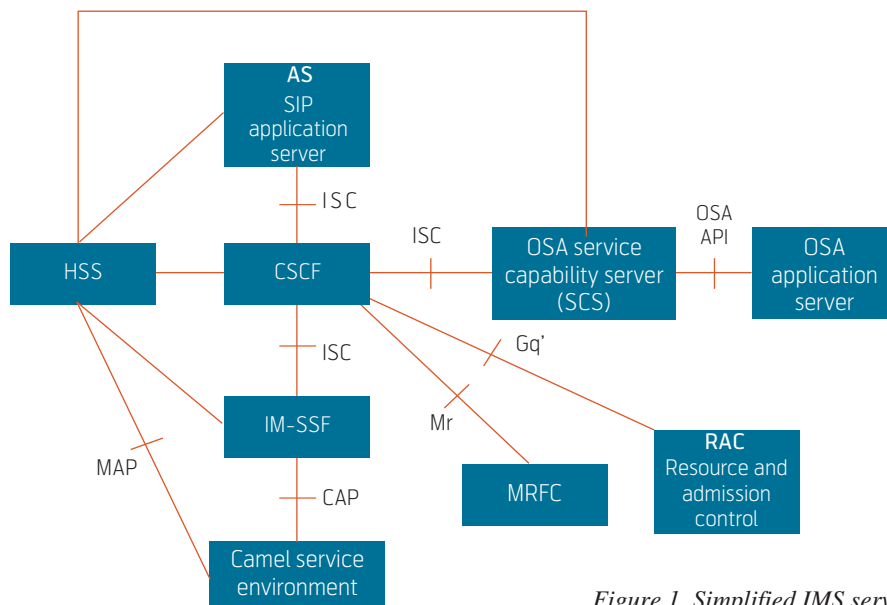


Figure 1 Simplified IMS service architecture

with session control for mobile terminals. The session control including management of dynamic inclusion/exclusion of elements in a session that is offered by SIP were important driving forces. In recent years, it has become ever clearer that IMS can be used as an overlay that enables access also via fixed lines. This has made IMS the tool devised e.g. by ETSI TISPAN for future access independent service offerings (i.e. for FMC). The service-oriented core of this IMS specification with TISPAN modifications [3] is depicted in Figure 1. This represents a functional expansion of the Core IMS subsystem as shown in Figure A1 of Appendix A.

IMS also provides the possibility to create more advanced services with the help of application servers. A third party operator, who only receives revenue from the use of a service, can in principle offer an application server.

The following describes the components of the IMS architecture as shown in Figure 1. The core components of this architecture as seen from a service perspective is the Call Session Control Function (CSCF) cooperating with Application Servers via the IP multimedia service control (ISC) interface. The ISC applies SIP signalling as defined in [4], [5] and [6], and the Application Servers can be:

- SIP Application Servers (AS) – which may host and execute services. The SIP Application Server may influence and impact on the SIP session on behalf of the end systems, depending on the services.
- IP Multimedia – Service Switching Function (IM-SSF) is an IN type of application server; the purpose of which is to host the Customised Applica-

tions for Mobile network Enhanced Logic (CAMEL) network features (i.e. trigger detection points, CAMEL Service Switching Finite State Machine, etc.) and to interface to Camel Application Part (CAP) as specified in 3GPP TS 29.078 [7].

- OSA service capability server (OSA SCS) – which offers access to the IMS for the OSA Application Server. This provides a standardized way for third party secure access to the IM subsystem.
- OSA Application Server – The OSA reference architecture defines an OSA Application Server as an entity that provides the service logic execution environment for client applications using the OSA API as specified in 3GPP TS 29.198 [8].

All the application servers (including the IM-SSF and the OSA SCS) behave as SIP application servers via the ISC interface. The Application Servers can also interact with the Multimedia Resources Function Controller (MRFC) via the S-CSCF (ISC and Mr interfaces). This enables application servers to control Multimedia Resource Function (MRF) processing (e.g. interactive voice response equipment).

2.1 Sessions and preconditions

Signal sequences for context activation, discovery, registration, call and session establishment are described in [5]. The following gives an overview of the signalling procedures required for session establishment. The main steps are:

1 System Acquisition and Attach:

After being powered up, the UE locks on to the cellular system. Once the appropriate cell is selected, the UE initiates the GPRS attach procedure. (These

steps are confined to the layer 1 and 2 signalling, and are not presented in further detail.)

2 Data Connection Set-up:

Once the UE is attached on layer 2, the next step is to establish the signalling Packet Data Protocol (PDP) data connection or “pipe” to the IMS (i.e. the signalling PDP context). The mobile terminal does not know the IP address of the P-CSCF at this point. The data connection is established by applying a PDP Context Activation message sequence. This creates the path required to carry SIP related signalling messages to the P-CSCF through the GGSN, which is the gateway to the P-CSCF (Figure 2). The response to the PDP Context Activation message includes the identity of the P-CSCF, and the PDP context will provide the UE with an IPv6 address to be used as the UE host address for the duration of the PDP signalling context. Both DHCP and DNS may be involved in this procedure. As the next step the UE carries out registration with the identified P-CSCF.

3 Registration:

Before engaging in an IP Multimedia session, the UE must perform the registration operation to let the IMS Core Network (CN) know the location of the UE. This registration is an application of SIP registration that opens for use of various SIP/IMS services. The UE acts as a SIP client and sends a SIP registration message (SIP REGISTER) to its home IMS system (i.e. to the S-CSCF) through the local P-CSCF.

4 Session Setup:

After a signalling PDP context is activated and Registration is completed, the mobile terminal can establish or accept an incoming session.

The main component of the IMS involved in SIP signalling, is the CSCF (Figure A1 in Appendix A and Figure 3). The CSCF SIP servers perform a number of functions such as multimedia session control and address translation function. In addition, the CSCF must manage service control, voice coder negotiation for audio communication, and Authentication Authorisation and Accounting (AAA). The CSCF plays three roles:

- The Proxy CSCF (P-CSCF) role,
- The Interrogating CSCF (I-CSCF) role, and
- The Serving CSCF (S-CSCF) role.

The S-CSCF is the key server with the main responsibility for the mobile’s session management and service provision.

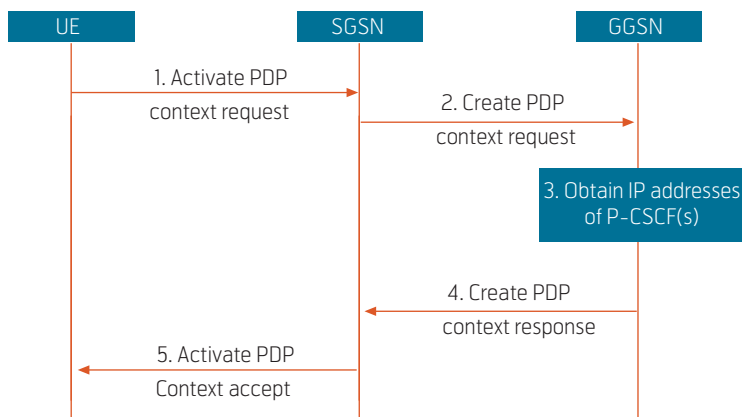


Figure 2 P-CSCF discovery using PDP Context Activation

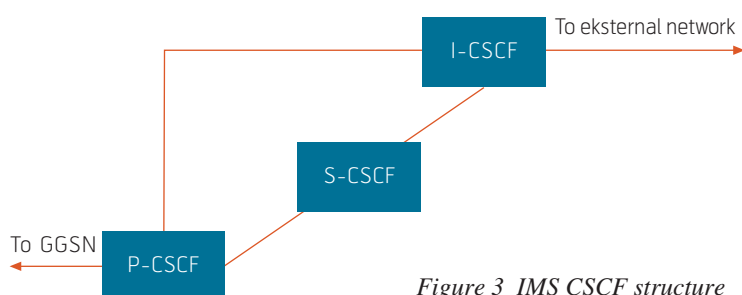


Figure 3 IMS CSCF structure

2.2 SIP methods and transactions

SIP [4] is the signalling protocol used between the Mobile Host (MH) or User Equipment (UE) and the IMS, as well as between the internal IMS components. SIP is used by IMS to offer such basic services as:

- User registration
- SIP based Sessions with the INVITE method
- Instant messaging
- Presence services
- Etc.

With the exception of the SIP ACK-message confirming the final response, there is always a dialog with Request/Response messages related to SIP transactions. Nested transactions are commonly used. The complete description of 3G SIP and the involved procedures can be found in [4], [5] and [6].

The final end-to-end signalling scenarios (after Registration) are composed of one signalling procedure from each of the following sets:

- Mobile originated signalling procedures (UE to SGSN#1)
- Serving-CSCF-to-Serving-CSCF procedures (SGSN#1 to SGSN#2)
- Mobile terminating procedures (SGSN#2 to UE).

Service Based Local Policy (SBLP) for QoS control for the UE, SGSN and GGSN is defined. This allows both the originating and terminating party to choose to use the QoS procedures on the local access. Both calling and called party establish a satisfactory PDP context (defining the QoS requirements) on their respective accesses. It is assumed that the core network is DiffServ enabled, and the combination of the QoS procedures in the access networks and the Diff-Serv enabled core network guarantees end-to-end quality of service under normal conditions. All signalling applies the signalling PDP context with its QoS definition, while a separate PDP context with the required QoS characteristics is established for the new session. This allows a SIP session to request the compliance to a set of QoS preconditions.

3 Fixed broadband IMS NGN

The Next Generation Network (NGN) is a network concept that aims at providing a framework to encompass the large variety of existing and emerging protocols and services, facilitate a further evolution of these, decouple this evolution from the underlying network infrastructure, and facilitate the interfacing of a plethora of available media. The rationale behind NGN lies well founded in paradigm shifts that have been taking place within the technological solutions and the business models in the telecom industry as a whole. NGN standards development takes place in all the major Standards Development Organisations (SDOs) including ITU-T, ETSI (TISPAN), 3GPP, and these SDOs are supported by the Internet Engineering Task Force (IETF) protocol development work.

The ETSI TISPAN [9] and the ITU-T are both involved in the Fixed Broadband IMS NGN standardization. The work takes the 3GPP Release 6 IMS specifications (completed as of now) as the basis. TISPAN will provide FMC requirements to the 3GPP for consideration for standardisation as part of Release 7. TISPAN Release 1 for the NGN was made available by the end of 2005.

The following reviews the architecture as defined for the TISPAN NGN, focusing on areas of importance for global VoIP provisioning. ETSI TR 180 001 [10] describes Release 1 of the TISPAN NGN, while the functional architecture is defined in [11]. The functional architecture shown in Figure 4 complies with the ITU-T general reference model for next generation networks and is structured according to a service layer and an IP-based transport layer. The service layer comprises the following components:

- Core IP Multimedia Subsystem (IMS) with SIP based session control,
- PSTN/ISDN Emulation Subsystem (PES),
- Other multimedia subsystems (e.g. streaming subsystem, content broadcasting subsystem ...) and applications,
- Common components (i.e. used by several subsystems) such as those required for accessing applications, charging functions, user profile management, security management, routing data bases (e.g. ENUM), etc.

The architecture allows addition of new subsystems over the time to cover new demands and service classes. Each subsystem is defined as a set of functional entities and related interfaces. As a result implementors may choose to combine functional entities where this makes sense in the context of the business models, services and capabilities being supported.

IP-connectivity is provided by the transport layer, under the control of the network attachment subsystem (NASS), and the resource and admission control subsystem (RACS). These subsystems hide the transport technology used in access and core networks below the IP layer, thus implementing a platform for realisation of access independence.

The functional entities that make up a subsystem may be distributed over network/service provider domains (e.g. the network attachment subsystem may be distributed between a visited and a home network). The following chapters describe the NGN subsystems in more detail starting with the IMS.

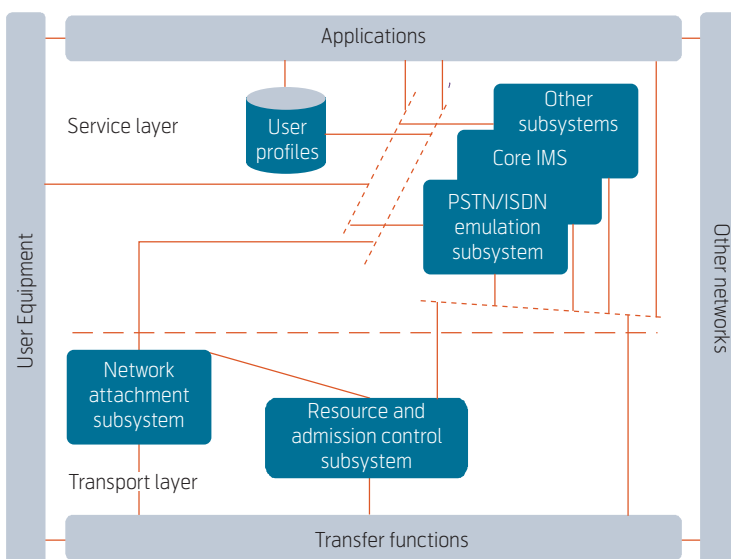


Figure 4 TISPAN NGN overall architecture

4 Services across multiple domains

The traditional and basic voice domains are the PLMN, PSTN and the ISDN. These domains provide ubiquitous voice services with well-defined quality based on circuit switching. All new voice services need interconnect with these basic voice services to provide value to their users. Figure 5 shows the interworking for VoIP. This includes the functionality required due to different standards for session control and bearer technology. The interworking shown in Figure 5 also includes cases not considered to be within the scope of 3G. In implementations of Interworking Units (IWU) it may be beneficial to combine or reuse functionality between different units. This aspect is not considered here. The following describes the (VoIP) scenarios of Figure 5:

- 1 The interconnect between the PSTN and PLMN PS domain voice services (VoIP) involves conversion between MAP and ISUP in the control plane, and transcoding and CS-PS bearer conversion at the user plane. The technical solutions involved are described in the “IMS VoIP interworking with PSTN/ISDN/CS domain” paragraph.
- 2 This is the traditional PSTN-PLMN interworking that takes place in the CS domains of both systems. It involves both signalling conversion and transcoding.
- 3 PLMN CS-PS interworking is similar to scenario 1 above, but the signalling for the 3G PS domain is based on SIP, so the interworking will involve SIP-ISUP interworking in the control plane. The “IMS VoIP interworking with PSTN/ISDN/CS domain” paragraph describes the technical solutions.
- 4 The interworking between the PLMN IMS/PS domain and the IP SIP domain is required to take care of security aspects and different SIP versions. The PLMN may demand to hide topology information from the open Internet. This may be handled by a topology-hiding inter-network gateway (THIG) located in the I-CSCF, at the PLMN border. The other aspect of this interworking is the handling of the difference between the IETF base SIP and the 3GPP SIP. This aspect of interworking is covered in the “Interworking between IETF core- and 3GPP SIP” paragraph.
- 5 Scenario 5 is required to allow interworking with the relatively large installed H.323 VoIP base.

6–8 These scenarios are already implemented and in operational use. They are therefore not further described.

- 9 Interworking between the PLMN CS domain and the H.323 VoIP service is required since certain 3G access networks can't sufficiently support conversational services like VoIP. The 3GPP has therefore established an architectural alternative for using existing CS bearers in association with an IM session. The solution involves adding one or more IP multimedia component(s) to a CS call to create a combinational service. Another reason for this interworking scenario is the 3GPP option of handing over a PS domain call (e.g. established at a WLAN hotspot) to the CS domain upon departure from the hotspot. An alternative that may be used is going through both IWU 3 and IWU 5. This may be feasible depending e.g. on the voice codecs applied.

An alternative to direct interworking with a domain is to go via domains already offering interconnects. This may be feasible in some cases, but repeated transcoding can introduce unacceptable delay for conversational services.

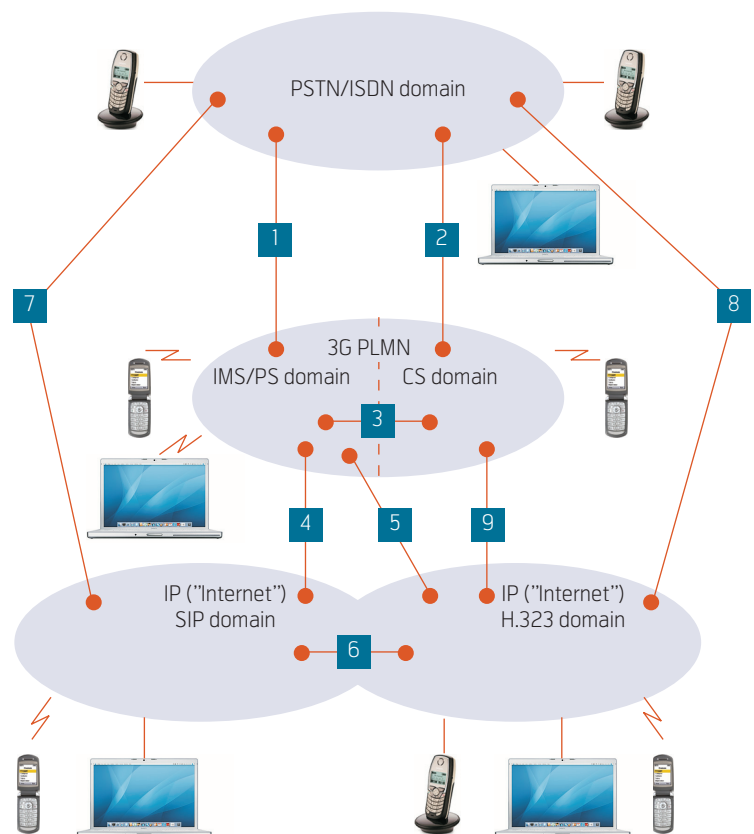


Figure 5 VoIP interconnect scenarios

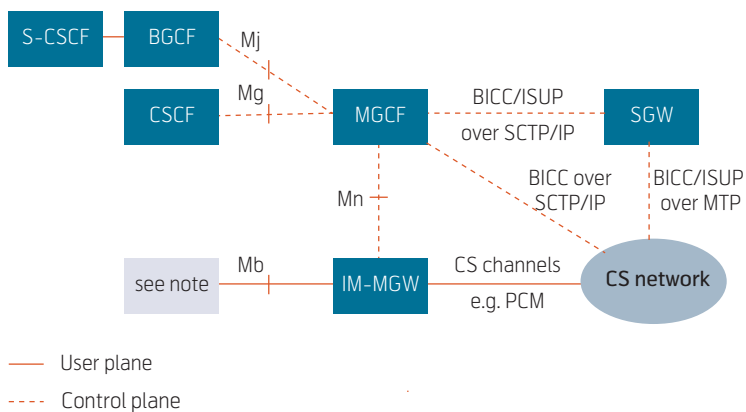


Figure 6 IM CN subsystem to CS network logical interworking reference model. Note: The IM-MGW may be connected to various network entities, such as a UE e.g. for voice (via a GTP Tunnel through a GGSN), an MRFP, or an application server

4.1 IMS VoIP interworking with PSTN/ISDN/CS domain

4.1.1 Architecture

The interworking reference model shown in Figure 6 is developed to support interworking for voice calls. This interworking between the IMS Packet Switched (PS) domain and the Circuit switched (CS) domain is defined in 3GPP TS 29.163 [12]. The CS domain is the ISDN/PSTN or the CS domain of the Public Land Mobile Network (PLMN). The MGCF performs SIP to BICC (BICC is the call control protocol used between nodes in a network that incorporates bearer

independent call control) or SIP to ISUP call related signalling interworking. In order to support existing network capabilities, it is required that IMS also supports endpoints (e.g. UE, MRFP, MGCF for interworking with the PSTN) able to send or receive DTMF tone indications using the bearer, i.e. inband signalling. An additional element for bearer interworking is the interworking of these DTMF tones and out-of-band signalling between one network and another. In such a case, the IM-MGW must provide tone generation and may provide detection under the control of the MGCF (Media Gateway Control Function).

The optional SGW (Signalling Gateway Function) performs conversion to or from BICC/ISUP based MTP transport networks to BICC/ISUP based SCTP/IP transport networks, and forwards the converted bearers carrying the signalling to or from the MGCF (Figure 6).

4.1.2 User plane and transcoding

The VoIP user plane protocol at the Mb reference point is IPv6 based and is defined by 3GPP TS 23.002. IPv6 and transport protocols such as RTP are used to transport user plane media packets to and from the IM CN subsystem entities like the UE or the MRFP. External legacy CS networks use circuit switched bearer channels like TDM circuits (e.g. 64 kbit/s PCM), ATM/AAL2(/1) circuit, or IP bearers to carry encoded voice frames. Voice bearers from the IM CN subsystem need to be connected with the bearers of other networks. Elements such as the IP Multimedia – Media Gateway Functions (IM-MGW), Figure 6, are provided to support such bearer interworking. One of the functions of the IM-MGW may be to support transcoding between a codec used by the UE in the IM CN subsystem and the codec being used in the network of the other party. The use of default codecs within the UE enables the IM CN subsystem to interwork with other networks either on an end-to-end basis or through transcoding. The traditional CS voice networks (e.g. PSTN, ISDN, CS domain of some PLMN) may be interworked for example by AMR to G.711 transcoding in the IM-MGW.

4.1.3 Voice features supported through interworking

The services that can be supported through the signalling interworking are the least common set supported by BICC or ISUP and SIP (Table 1). The MGCF will originate and/or terminate services or capabilities that do not interwork across domains. Table 1 lists the services that are seamlessly interworked for voice [12].

Service
Speech / 3.1 kHz audio
En bloc address signalling
Overlap address signalling from the CS side towards the IMS
Out of band transport of DTMF tones and information (BICC only)
Inband transport of DTMF tones and information (BICC and ISUP)
Direct-Dialling-In (DDI)
Multiple Subscriber Number (MSN)
Calling Line Identification Presentation (CLIP)
Calling Line Identification Restriction (CLIR)
Connected line presentation (COLP)
Connected line restriction (COLR)

Table 1 Interworking capabilities between BICC/ISUP and SIP profile for 3GPP

4.2 Interworking between IETF core- and 3GPP SIP

Interoperability between the IMS and the Internet is required to allow multimedia sessions between the Internet users and the IMS users. There are two ways to achieve this interoperability. The preferred way is to maintain compatibility between the SIP signalling procedures applied for the Internet and the PLMN. This paragraph discusses the challenges of such SIP-to-SIP interworking.

The 3GPP IMS definition 3GPP TS 23.228 [2] states that depending on operator policy, the S-CSCF may forward the SIP request or response to another SIP server located within an ISP domain outside of the IM CN subsystem. It is possible that the external SIP client does not support one or more of the SIP extensions defined for IMS end points (e.g. Preconditions, Update, 100 Rel) as described in 3GPP TS 24.229 [6]. In such cases the UE or other SIP user agents within the IMS should be able to fall back to basic SIP procedures which allow interworking towards the external client.

The main technical issues related to SIP Interoperability are:

- Can the already available basic IETF SIP user agents on the Internet be used to access the IMS services?
- Can the already available SIP application servers on the Internet be used as application servers for the IMS?
- Can SIP signalling openly be exchanged between the IMS and the Internet?

The IETF and 3GPP are working in close collaboration in order to achieve the required interoperability. However, IMS and the Internet are partially competing against each other and thus the actual barriers of interoperability might turn out to be non-technical.

Anyhow, ETSI TISPAN [9] is considering these interworking aspects as a part of the TISPAN requirement work. This will hopefully ensure interoperability limited only by involved security requirements.

4.2.1 3GPP SIP extensions

The SIP protocol is designed for extensibility. Extensions may define new methods and header fields at any time. A SIP network element must therefore not refuse to proxy a request because it contains a method or header field it does not recognize. This implies that the interoperability between 3GPP and IETF SIP in principle largely depends on the User Agents (UAs)

in the involved end systems. However, the 3GPP SIP may have defined extensions to the proxy functionality that may be missing in the Internet, thereby effectively causing a need for limiting the functionality even when the end-systems on the Internet have implemented the 3GPP UA. This means that, conditional to the availability of the 3GPP extensions in the UA, all end systems on the Internet in principle can interoperate with the 3GPP functionality only dependent on the required 3GPP proxy processing. However, the 3GPP SIP extensions are based on certain assumptions regarding network topology, linkage between SIP and lower layers, and the availability of transitive trust. These assumptions are thus generally not applicable in the Internet as a whole. The expected outcome is that the 3GPP SIP proxy connecting to the Internet (i.e. the I-CSCF(THIG)) will need to filter out information that can not be submitted openly on the Internet. Alternatively, directed security measures could be enforced e.g. to allow certain users or user groups to receive this information also on the Internet.

4.2.2 Harmonisation between 3GPP and IETF SIP

3GPP notified the IETF SIP and SIPPING working groups that existing SIP documents provided almost all the functionality needed to satisfy the requirements of the IMS, but that they required some additional functionality in order to use SIP for its purpose. These requirements are documented in an Internet Draft [13] that was submitted to the SIPPING Working Group. Some of these requirements are satisfied by IETF chartered extensions, while other requirements were applicable to SIP, but not sufficiently general for the SIP Working Group to adopt for the IETF baseline SIP. The current state of harmonization is described in 3GPP TS 24.229 [6], IP Multimedia Call Control Protocol. On the subjects of SIP methods, SIP headers, option tags, status codes and session description types, there are no definitions within the 3GPP technical specification over and above those defined in the referenced IETF specifications.

The following private SIP header extensions are however defined for 3GPP applications, [6] and [14]:

- Extension to WWW-authenticate header
- Extension to Authorization header
- Tokenized-by parameter definition (various headers)
- P-Charging-Vector header
- Orig parameter definition.

Interoperability is however helped by the fact that the relaying network elements do not have to support all SIP (header) extensions:

- User agents can negotiate the required and supported SIP extensions.
- If the network elements between the user agents receive extended SIP messages, which they can't understand, they simply forward the message as it was received (unless filtering is required).
- It is up to the receiving party to fulfil the functionality of extensions.

Intelligence may thereby be concentrated on the edges of the network.

The Annex A of the 3GPP TS 24.229 [6] specifies the profiles of IETF RFCs for 3GPP usage. This profile specification helps ensure implementation of compatible 3GPP SIP versions.

5 Standards Development Organizations and NGN

5.1 IMS, 3GPP and ETSI TISPAN

ETSI TISPAN [9] has been studying material available in 3GPP for reuse in the NGN specification. There is a need for material specifying the architecture (functional entities and interfaces) of IMS in a manner independent of the rest of the 3GPP architecture. The TISPAN NGN architecture is currently specified in 3GPP TS 23.002 embedded with the rest of the material for the 3GPP architecture. As such it is difficult to use this approach for the evolving NGN architecture. Therefore TISPAN has currently taken the approach of a totally separate and independent specification to specify this material. It is understood that this has also been an issue within the equivalent ITU-T NGN work.

ETSI TISPAN only wants to include the issues needed to realize the targeted NGN system. This issue of reuse is not streamlined between the two organizations, and the current structure of work coming from 3GPP does not appear to lend itself to reuse within the context of a generic access technology (i.e. TISPAN wants access independent specifications). ETSI TISPAN would like to restructure this work in order to allow common specification for use by both 3GPP IMS and NGN IMS. As an example the ETSI TISPAN is willing to reuse the multimedia functionality of the HSS. However, the HSS as currently described in 3GPP TS 23.002 is difficult to represent in the context of the ETSI TISPAN architecture, as it encompasses HLR/AUC functionality that is not relevant to the IP-networks being defined within the TISPAN architecture. Therefore TISPAN would like to assign a component name to the multimedia func-

tionality of the HSS, excluding HLR/AUC, which could be reused as a functional entity name in the TISPAN architecture. This is an example of undesirable implementation assumptions taken in the 3GPP specifications.

5.2 3GPP All IP Network (AIPN) with New Access Technology

The general feeling of the 3GPP SA2 architectural working group is that there is very much to do to upgrade and enhance the 3GPP architecture to better support the future market. The 3GPP is now focusing on the future 3G-system evolution. The aim of this work is to construct a new IP based mobile system taking the 3GPP system as a starting point and to modify (migrate/evolve) it as to make it less complex and targeted to the future communication demands. Important parts of such a long-term evolution include reduced latency, higher user data rates, improved system capacity and coverage, and reduced overall cost for the operator. Additionally, it is expected that IP based 3GPP services will be provided through various access technologies including various types of radio systems and fixed access systems. Support of seamless mobility between heterogeneous access networks is a highly stated requirement. The target deployment for the 3GPP NGN is in the 2010s, and the 3GPP is expecting the standardization work to be completed by mid-2007. The all-IP core network is expected in the same timeframe as the evolved UTRAN (E-UTRAN).

The IETF netlmm working group is aiming at meeting the needs for localized mobility management for the above NGN development.

5.2.1 List of issues

The following is an architectural level list of selected key questions related to the new system architecture and future releases:

- How to achieve mobility within the new Access System?
- How to add support for non-3GPP access systems?
- Inter-access-system mobility?
- Is user access control/authentication per access system or more centralized for multiple access systems?
- In case a UE accesses/attaches multiple Access Systems in parallel: how does reservation of guaranteed resources work? Are multiple reservations in parallel required (same resource on every Access System) to allow for fast change between Access

Systems? Or, does a mobility/handover mechanism reserve resources during the mobility/handover process?

- Shall inter Access System mechanisms and signaling for load sharing and mobility be generic for all Access Systems, or peer-to-peer between interoperable Access Systems?
- Will the Access Systems have an idle or paging mode? And, shall the wake-up work over multiple Access Systems (e.g. paging in multiple Access Systems in parallel)?
- May functions be transferred to application/services level (e.g. mobility supported by IMS services)? If yes, to which extent is this feasible for application/services?
- How is data compression provided for the different access systems?
- IMS local services: This allows the support of local services when roaming. This has been a difficult topic since the S-CSCF is always in the home network.

work. This was initially part of R6 3GPP specifications, but is now part of R7.

- Emergency calls: Support of emergency services is required to enable migration to all-IP.

5.3 Combining CS and PS domain

5.3.1 Combinational services

Some 3G access networks do not have the required PS domain efficiency for acceptable support of conversational services like VoIP. The 3GPP study [15] therefore investigates architectural requirements and architectural alternatives for using existing CS bearers in association with an IM session. The solution advocated is to add one or more IP multimedia component(s) to a CS call to create a combinational service. The CS and IMS components are established between the same participants.

Combinational services are also intended to enable a user participating in a mobile-to-mobile Circuit Switched (CS) conversation with another user, e.g. to take a picture/video with the built in camera in the mobile and transmit the picture to the other party during this conversation.

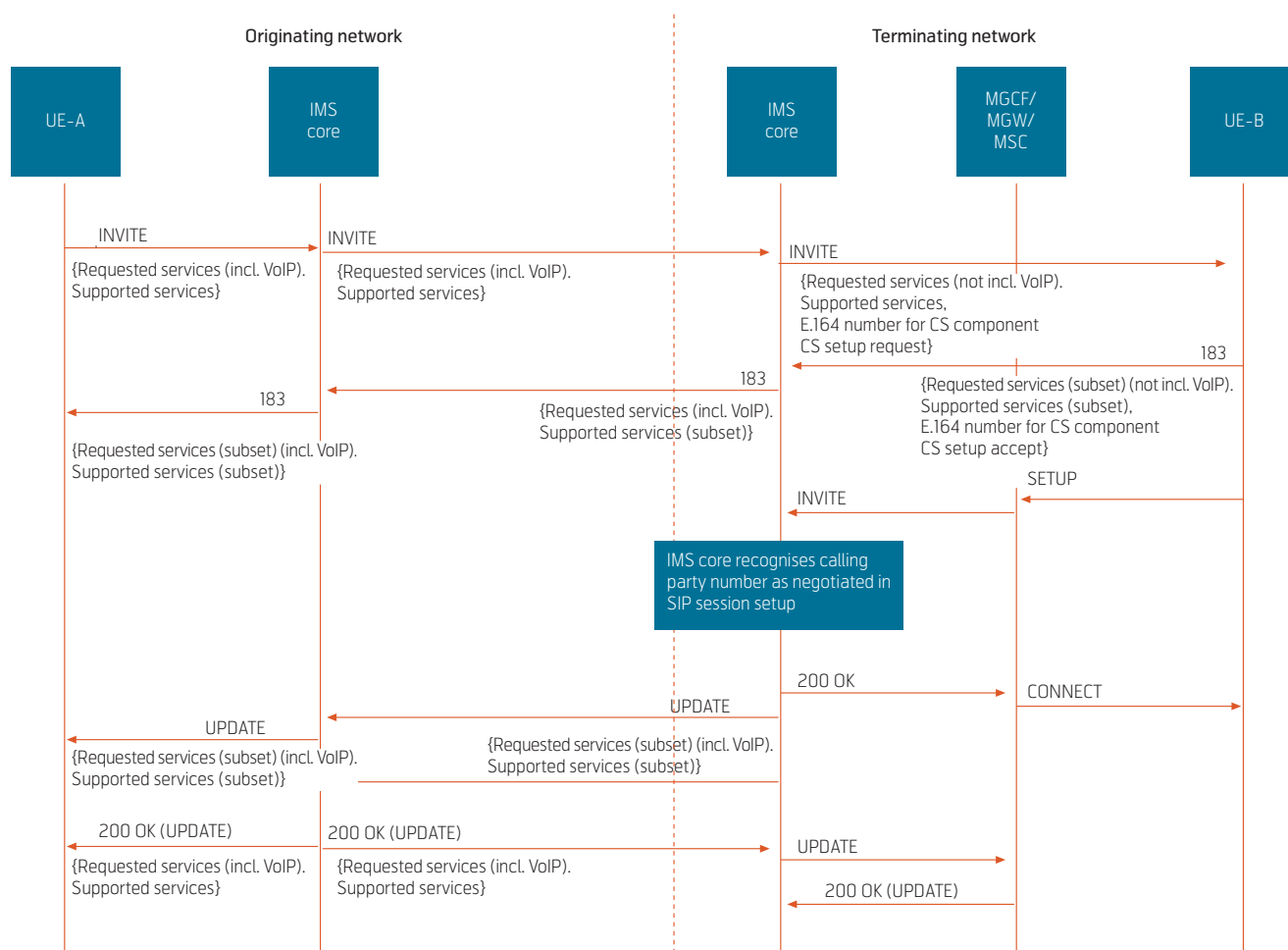


Figure 7 Call from pure VoIP endpoint using client-to-network CS call

The study [15] considers requirements for evaluation of potential architectural solutions and proposes continuing development of a converged architecture. It is recommended [16] to develop the new technical specification for combinational services as a part of 3GPP Release 7 specifications. The target is to arrive at an architectural solution that is completely transparent for the end-user, and is easily interoperable with existing IMS services and other networks that do not use this solution.

Figure 7 shows the first part of a signal sequence, defined for the converged architecture [15], for a call originating from a VoIP only handset terminating (at UE-B) in the CS domain. This flow uses an E.164 number negotiation to inform the terminating UE of the calling line identity that will identify a forthcoming incoming CS call and for the terminating UE to indicate its willingness to accept such a call. This CS call will be interworked to the VoIP component from the pure VoIP endpoint.

This flow uses E.164 number negotiation. It uses an additional indication to request the terminating UE to establish a CS call to the provided number. This call will then be interworked to the VoIP component at the pure VoIP endpoint. Alternatively, the CS call may be established first, with the IMS session being established after the CS call is in the Alerting state.

As the new mechanism proposed for E.164 number negotiation, for the possible addition of a CS call later in the session, requires new functionality either on the SIP or the SDP protocol level, it is assumed that this functionality of delivering multi-component IMS sessions from pure VoIP parties to Circuit Switched Bearer (CSB) parties will not be available in the initial phase of combinational services launch.

5.3.2 PS handover to CS domain

IMS handover to CS domain allows an IMS call initiated e.g. on a WLAN through IMS to be handed over to a CS bearer when for instance losing WLAN coverage in the case when the GSM/GPRS network does not support real time QoS via IMS.

5.4 End-to-end QoS

The work on end-to-end QoS has created much debate within the 3GPP. The reason is that 3GPP has started work on inter domain QoS signaling to support roaming users PS based traffic. The GSM Association (GSMA) has published a document where they clearly state that DiffServ is the QoS mechanisms to be used across operator domains to preserve end-to-end QoS, and they claim that this will be the mechanism for many years ahead.

6 Selecting the NGN target for future development

6.1 General

The interconnection scenarios, shown in Figure 5, clearly indicate that there are some challenges related to the evolution towards a harmonised NGN. The current situation is characterised by a plethora of multimedia platforms (e.g. for VoIP production), and no common target for architectural development. This is as seen resulting in fragmentation and the need for extensive interworking. The consequence of this lack of technical harmonization is loss of advantage of scale, reduced value of new multimedia services caused by lack of ubiquity, and high complexity in networks and their operation.

Previously the International Telecommunications Union (ITU) was able to manage the required specification and harmonisation work to allow nations and operators to maintain the desired interoperability. However, today there is no such strong coordination, and much of the development is left to completely obey the laws of commercial competition. This is good in many respects, but without a common architecture setting the direction of development, and ensuring service interoperability and ubiquity, the users of telecommunication services and the operators will mainly suffer because of limitations in service availability, in quality, in traffic volume, and in incurred costs of service production.

6.2 Generic core architecture

Figure 8 sketches the top level of a proposed NGN service architecture that could benefit both operators and their subscribers. This architecture may be adopted to ensure both service ubiquity and network harmonisation.

The only interworking shown explicitly is towards PSTN / ISDN (1) service network, described in the "IMS VoIP interworking with PSTN/ISDN/CS domain" section, and towards the H.323 domain of the Internet (2) [17], [18]. All other interworking is implicit in the Figure 8, and characterised by all being against the common core effectively eliminating the need for the N square interworking arrangements for N different technologies (e.g. service networks).

The Application Server (AS) (Figure 8) is attached to the common core. Thereby being capable of supplying ubiquitous access independent services in the core representation. These services will be interworked with the service representations at the existing service specific networks at the edge of the common core.

There will be ample room for competition within the framework despite the harmonised core that will effectively set the direction of evolution. Vendors can gain from increase in scale in core technologies. The number of different network elements will be reduced, and the competitive arena will move in the direction of new services, maintaining bridging of new and existing technologies.

The challenge is to agree on the core technologies. This is because of the huge invested interest in the different existing technologies allowing little room for selecting one as the core and basis for future development. The principal candidate technologies are the 3G cellular technologies and the IETF defined Internet technology. A good solution would probably be to select the best of breed from these candidates and to agree the following key characteristics and solutions for the new core architecture:

- Multimedia service architecture (e.g. SIP)
- Mobility management scheme for global (macro) mobility
- Core set of codecs
- IP bearer service with QoS control mechanism (e.g. DiffServ)
- Core VPN solution
- Common multicast mechanisms and protocols
- Naming and addressing without overloading the semantics of IP addresses
- Core Authentication Authorisation and Accounting (AAA)
- Core API for bearer access (enabling application portability).

A discussion of these points applied as components in a service-oriented architecture can be found in [19].

One important point is that most of the above functionality is readily available. The major challenge will therefore be to agree on a choice for the target architecture. Agreeing on such a selection together with an open approach can be expected to create a very innovative and competitive market for global service production, benefitting the end-users, operators and vendors by increasing the overall revenue in telecommunications.

6.3 IETF core architecture

The IETF based alternative core approach is shown in Figure 9. Here Mobile IP is chosen for global mobility management (i.e. macro mobility). The UMTS is merely considered to be an access technology, on the level of a wireless local area network, with a MAP based local mobility management.

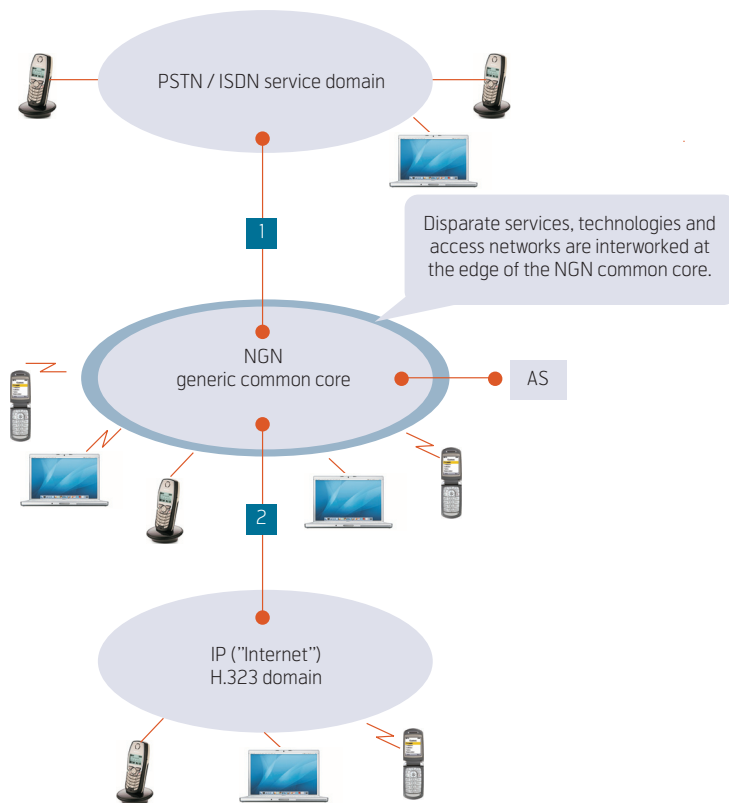


Figure 8 Target core for service innovation

6.4 3G GRX-based core architecture

A competing alternative for the core is the GPRS/GRX backbone as shown in Figure 9 and expanded in Figure 10. (The GPRS architecture with IMS is briefly described in Appendix A.) The GPRS Roaming Exchange (GRX) [20] is built on a private or public IP backbone and transports GPRS roaming traffic via the GPRS Tunnelling Protocol (GTP) [21] between the visited and the home PLMN (Public Land Mobile Network). A GRX Service Provider (operator) provides of a set of routers with links connecting to the GPRS networks and links connecting to other GRX nodes for peering. The GRX service provider thereby acts as a hub. There is no need for a GPRS operator to establish a dedicated connection to each roaming partner; instead the GPRS operator establishes a connection to the GRX. This allows easier implementation of new roaming relations and rapid time to market for new operators. An alternative is therefore to apply an enhanced GRX backbone as the common converged core.

When the GGSN (Gateway GPRS Support Node) and the SGSN (Serving GPRS Support Node) are located in different networks, they may be interconnected via the IP based Gp interface. This interface provides similar functionality to that of the Gn interface (shown in Figure A1 in Appendix A), however it usually includes extra security functionality based on

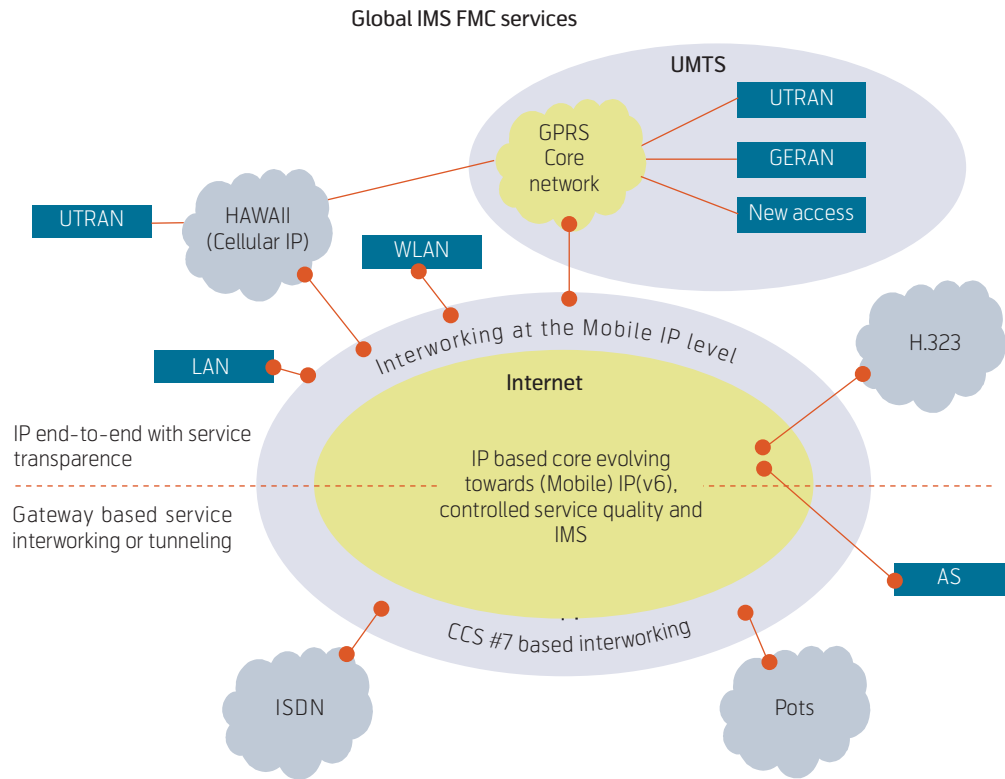


Figure 9 IETF based core alternative

mutual agreements between operators. The Gp Interface connects PLMNs together. Gp must support appropriate routing and security protocols to enable an end user to access its home services from any of its home PLMN's roaming partners. Many GPRS

operators/carriers have abstracted these functions through the GRX (Figure 10). This function is typically provided by a 3rd party IP network offering VPN (Virtual Private Network) service that connects all the roaming partners together. The GRX service

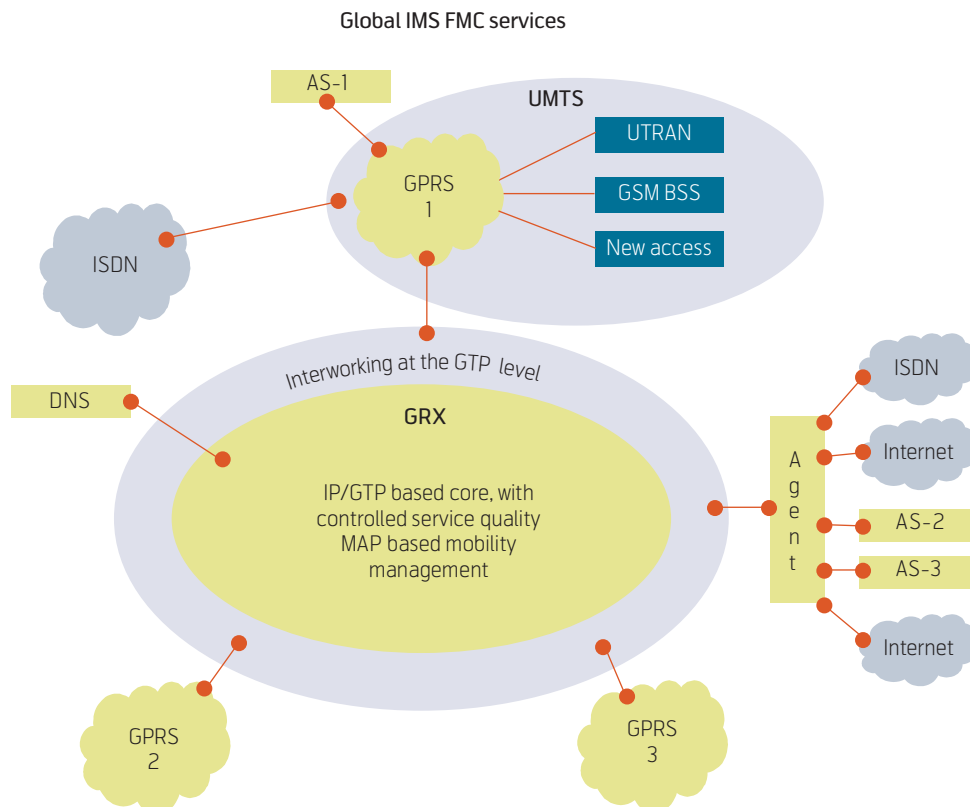


Figure 10 GRX/GPRS based logical core alternative

provider handles aspects of routing and security between the GPRS networks. The GRX may additionally operate DNS and ENUM services.

The role of the Agent (Figure 10) is to facilitate interworking and interoperability from both a technical and commercial perspective, key functions include:

- Contract brokering – access to multiple operators/providers via a single agreement;
- Service control and interworking (e.g. for the Internet);
- Accounting and settlement – a single partner for inter-network accounting;
- Mobility management for end systems roaming outside the 3G-domain (e.g. GTP/MAP and Mobile IP interworking);
- IP packet transport (control plane and user plane) – with the prerequisite QoS.

The GRX specifications [20] now state that the GRX shall also support the conversational class of services:

- Max Delay 20 ms
- Max Jitter 5 ms
- Packet Loss 0.5 %
- SDU Error Ratio 10^{-6}

This makes the GRX backbone well suited for e.g. VoIP and conversational applications.

The Agent depicted in Figure 10 is purely logical, and does not indicate relaying of user plane traffic at a central point. GPRS network 2 and 3 are shown without interworking and service infrastructure. These networks apply the interworking infrastructure and service infrastructure as offered by the Agent. (The Agent represents the home network for subscribers of GPRS 2 and GPRS 3.)

This hosting offered by the Agent may be utilised as follows:

- Common service production for all operators;
- Common service production for all affiliates of e.g. a multinational operator;
- Hosting of services for third party service providers.

In the basic GRX network only SGSNs and GGSNs implement the GTP protocol. No other systems need to be aware of GTP. The implications of interworking at the GTP level (i.e. for the Agent functionality)

needs further investigation, but is outside the scope of this presentation.

Subscriber identification is of significant commercial importance, and is a key aspect of the architecture. Operators of GSM/GPRS based cellular networks are mostly in favour of applying the SIM-card for this purpose. However, alternatives are needed as FMC and VoIP also require non-SIM capable user equipment access to the common core as a minimum for use of the basic voice service.

6.5 Administrative operations support systems

It is claimed that one of the most important aspects of the proposed harmonization, by introducing the IMS service overlay, is to give operators the opportunity to clean up and take control over the plethora of existing and partly incompatible administrative operations support systems. This has the potential of significantly reducing the operational and capital expenditure. How this is to be achieved in practice needs further study. However, introduction of a new set of administrative support systems integrated with the IMS, spanning the functionality ranging from CRM, via Order to Billing is an interesting but ambitious option. Such a new system might prove commercially viable if based on common harmonised data definitions.

7 Conclusions

Interoperability between the different voice services has to be ensured in order to maximize the value of the next generation IP based voice and multimedia services both for the end users, and for the industry as a whole. This presentation has shown that this interworking for 3GPP based systems is exceedingly complex. It is therefore proposed to define a new architecture for the core network. This core architecture will both simplify the interworking, and act as target architecture for future development. However, it can be difficult to standardise this architecture because of conflict of interests between the involved industry players. Operators may therefore protect their interests by choosing their internal core, and benefit from reduced complexity and minimising duplication of network and service development. The technologies not selected as core may be phased out as the core implements the capabilities and capacity to take over the service production. IMS for Fixed Mobile Convergence is the first implementation step, but there are still multiple issues, like target topology, naming and addressing, global and localized mobility management etc., that are to be clarified before the full core network can be implemented. The proposed core architecture is also a tool for migration, and in princi-

ple allows adoption of the best solutions from each of the existing mobile and fixed technologies for the core. This is possible while still maintaining the opportunity and flexibility for innovation in services and in access technologies.

The ETSI TISPAN work can be viewed as a first step by standardising the common core overlay technology. However, as explained, much more than an FMC version of the IMS needs to be settled. The conflicts lie in the zone between the 3G and the IETF solutions. This represents a major challenge when it comes to selecting the core. Uniting the forces to arrive at a common core by selecting the best alternatives would in the long run benefit most parties.

The 3GPP has now started the work to construct a new IP based mobile system (AIP NGN) taking the 3GPP system as a starting point. Modifications are planned in order to arrive at a less complex system targeted to the future communication demands. Support of seamless mobility between heterogeneous access networks is a highly stated requirement. The 3GPP is expecting the standardization work to be completed by mid-2007. The new recommendations that will be produced will hopefully realize a harmonized architecture with a common core as suggested here.

The proposed core with an IMS overlay is anticipated to give operators the opportunity to clean up and take control over the plethora of administrative operations support systems thereby reducing the operational and capital expenditure. However, this needs further investigation.

8 References

- 1 3GPP. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; General Packet Radio Service (GPRS); Service description; Stage 2 (Release 6)*. December 2004. (TS 23.060 V6.7.0)
- 2 3GPP. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS); Stage 2 (Release 6)*. December 2004. (TS 23.228 V6.8.0)
- 3 ETSI. *Telecommunications and Internet Converged Services and Protocols for Advanced Networking (TISPAN); NGN Release 1; NGN-IMS Stage 2 definition (endorsement of TS.23.228)*. September 2005. (TS 182 006 V 0.0.7)
- 4 Rosenberg, J et al. *SIP: Session Initiation Protocol*. June 2002. (RFC 3261) URL: <http://www.ietf.org/rfc/rfc3261.txt?number=3261>.
- 5 3GPP. *3rd Generation Partnership Project; Technical Specification Group Core Network; Signalling flows for the IP multimedia call control based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 5)*. January 2005. (TS 24.228 V5.11.0)
- 6 3GPP. *3rd Generation Partnership Project; Technical Specification Group Core Network; IP Multimedia Call Control Protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3 (Release 6)*. January 2005. (TS 24.229 V6.5.1)
- 7 3GPP. *3rd Generation Partnership Project; Technical Specification Group Core Network; Customised Applications for Mobile network Enhanced Logic (CAMEL) Phase 4; CAMEL Application Part (CAP) specification (Release 6)*. December 2004. (TS 29.078 6.4.0)
- 8 3GPP. *3rd Generation Partnership Project; Technical Specification Group Core Network; Open Service Access (OSA); Application Programming Interface (API); Part 1: Overview (Release 6)*. December 2004. (TS 29.198-1 V6.3.1)
- 9 ETSI TISPAN homepage. URL: http://portal.etsi.org/Portal_Common/home.asp.
- 10 ETSI. *Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Release 1: Release Definition*. June 2005. (TR 180 001 V0.4.2)
- 11 ETSI. *Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN); NGN Functional Architecture Release 1*. August 2005. (ES 282 001 V 1.1.1)
- 12 3GPP. *3rd Generation Partnership Project; Technical Specification Group Core Network; Interworking between the IP Multimedia (IM) Core Network (CN) subsystem and Circuit Switched (CS) networks (Release 6)*. March 2005. (TS 29.163 V6.6.0)
- 13 Garcia-Martin, M. *3rd-Generation Partnership Project (3GPP) – Release 5 requirements on the Session Initiation Protocol (SIP)*. IETF Internet Draft, Work in Progress, October 11, 2002, Informational RFC, RFC 4083, on 2005-5-25. URL: <http://www.ietf.org/rfc/rfc4083.txt>.

- 14 Garcia-Martin, M et al. *Private Header (P-Header) – Extensions to the Session Initiation Protocol (SIP) for the 3rd-Generation Partnership Project (3GPP)*. January 2003. (RFC 3455) URL: <http://www.faqs.org/rfcs/rfc3455.html>.
- 15 3GPP. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Report on alternative architectures for combining CS Bearers with IMS; Release 6*. June 2005. (TR 23.899 V1.2.0)
- 16 3GPP. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Feasibility study on combined Circuit Switched (CS) calls and IP Multimedia Subsystem (IMS) sessions (Release 7)*. March 2005. (TR 22.979 V7.0.0)
- 17 Schulzrinne, H, Agboh, C. *Session Initiation Protocol (SIP)-H.323 Interworking Requirements*. July 2005. (RFC 4123) URL: <http://www.ietf.org/rfc/rfc4123.txt>
- 18 Singh, K, Schulzrinne, H. *Interworking Between SIP/SDP and H.323. Proceedings of the 1st IP-Telephony Workshop (IPTel'2000)*, Berlin, April 2000. [online] <http://www1.cs.columbia.edu/~kns10/publication/iptel2000.pdf>.
- 19 Grønbaek, I. *A Service Oriented Architecture for Multi-domain Mobility. Proceedings of ICT 2005 – 12th International Conference on Telecommunications*, Cape Town, South Africa, May 3–6, 2005. (ISBN 0-9584901-3-9)
- 20 GSM Association. *Permanent Reference Document IR.34, Inter-PLMN Backbone Guidelines, Version 3.5.2*. August 2004.
- 21 3GPP. *3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; General Packet Radio Service (GPRS); GPRS Tunneling Protocol (GTP) across the Gn and Gp interface (Release 6)*. June 2005. (TS 29.060 V6.9.0)

Appendix A: 3GPP reference architecture

The 3GPP reference architecture [1] with SIP interfaces and service control is shown in Figure A1. The following gives a brief sketch of the functional components and their interrelationships. This part is included to show how the VoIP and multimedia functionality fits together in the context of 3G.

Functional components

BGCF (Breakout Gateway Control Function) selects the network to which breakout of a call or session is to occur (e.g. to PSTN or other network).

- CSCF (Call Session Control Function) SIP based control function composed of Proxy CSCF (P-CSCF), Serving CSCF (S-CSCF) and Interrogating CSCF (I-CSCF). These functions are implemented as SIP servers and can incorporate the following functionality:
 - Proxy server is a server application that handles and may forward SIP requests;
 - Redirect server is a server application that redirects SIP requests;
 - Registrar is a server that only accepts REGISTER requests (typically co-located with a proxy or redirect server).
- DNS (Domain Name System – not shown in the figure) is used e.g. by the CSCF to resolve the IP addresses of IMS entities; E.164 numbers (ENUM DNS) and domain names.
- Customised Applications for Mobile network Enhanced Logic (CAMEL) see IM-SSF below.
- HSS (Home Subscriber Server) is responsible for storing user-related information: user identifiers, numbers, addresses, security policies, location, and user and service profiles. HSS includes all legacy HLR interfaces and functions. (In short: HSS = HLR + AAA)
- I-CSCF (Interrogating-CSCF) is the contact point with external networks. Also during registration or session establishment I-CSCF is used to find the correct S-CSCF applying information from the HSS. A THIG (Topology Hiding Inter-network Gateway) function within I-CSCF may be optionally utilized to hide the internal structure of operator's IMS network. The I-CSCF (internal THIG) functionality shall make it possible to restrict the following information from being passed outside of an operator's network: exact number of S-CSCFs, capabilities of S-CSCFs, or capacity of the network.
- IM-MGW (IP Multimedia – Media Gateway) IM-MGW is used for PSTN or CS-domain interconnection. Is controlled by the MGCF.
- IP Multimedia – Service Switching Function (IM-SSF) is an IN type of application server the purpose of which is to host the Customised Applications for Mobile network Enhanced Logic (CAMEL) network features (i.e. trigger detection points, CAMEL Service Switching Finite State

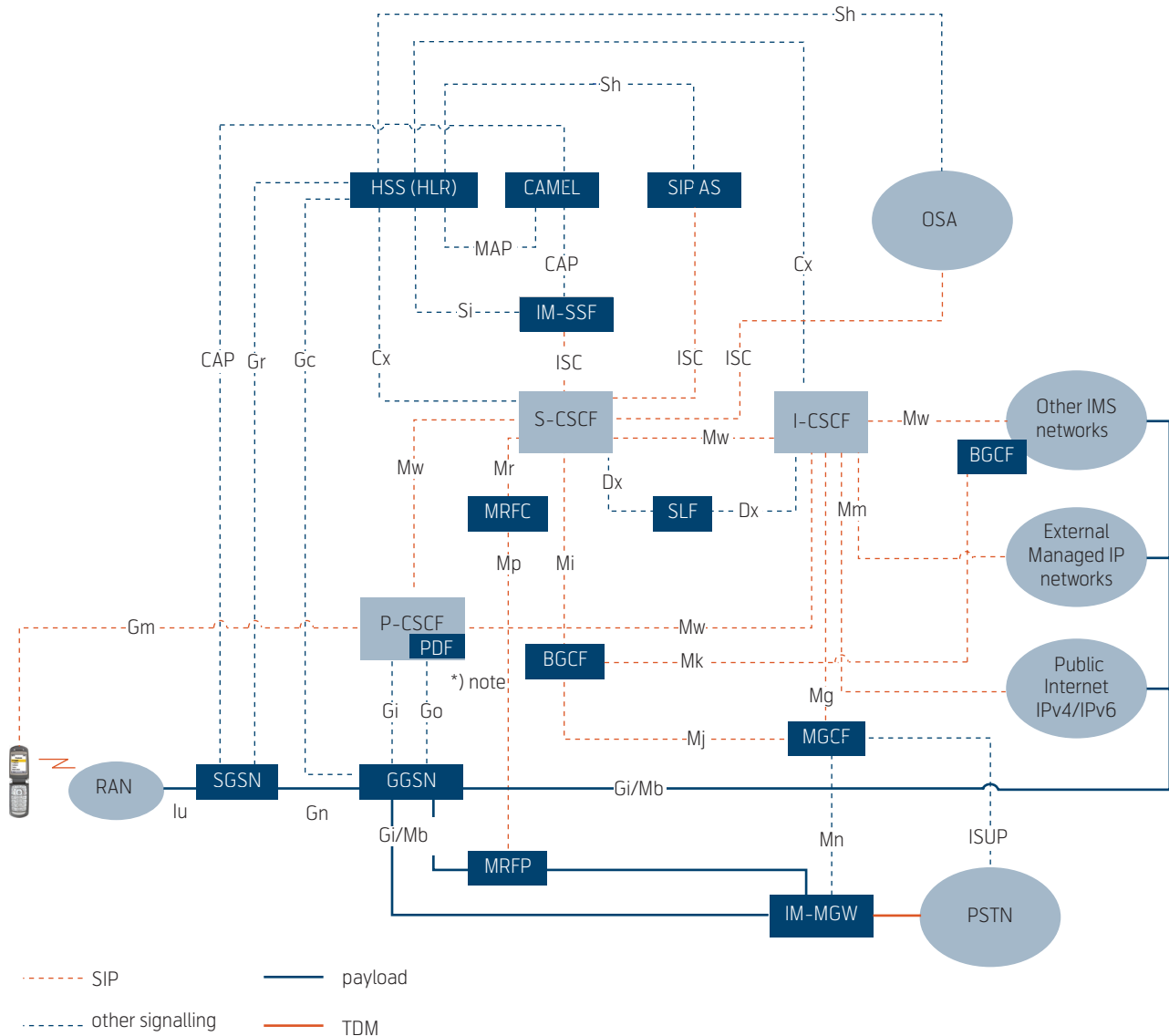


Figure A1 3GPP architecture with SIP and service control (IMS)

*) note: The PDF may not be colocated with the P-CSCF

- Machine, etc.) and to interface to Camel Application Part (CAP) as specified in 3GPP TS 29.078 [7].
- MGCF (Media Gateway Control Function) is used to control IM-MGW, and it converts legacy signalling protocols to SIP signalling.
- MRF (Multimedia Resource Function) performs multiparty call and multimedia conferencing functions. The MRF consists of two parts MRFC (Multimedia Resource Function Controller) and MRFP (Multimedia Resource Function Processor).
- OSA service capability server (OSA SCS) offers access to the IMS for the OSA Application Server. This provides a standardized way for third party secure access to the IM subsystem.
- P-CSCF (Proxy-CSCF) is the first contact point for the User Equipment (UE) within the IMS. The Policy Decision Function (PDF) is a logical entity of the P-CSCF, which manages resource control for the GGSN. The main function performed by the P-CSCF is forwarding of the SIP messages from the UE to other CSCFs and vice versa.
- S-CSCF (Serving-CSCF) performs the session and services control for the end point. The main functions performed by S-CSCF are registration, session control for registered end points and interaction with service platforms (e.g application servers). S-CSCF is always situated in the home network, even in the roaming situation and when the services of the visited network are used. Thus the subscriber is always registered by S-CSCF of the home network for the purpose of charging and HSS access.

- SIP Application Servers (AS) may host and execute services. The SIP Application Server may influence and impact on the SIP session on behalf of the end systems, depending on the services.
- SLF (Subscription Locator Function) is used to find the serving HSS. The SLF is not required in a single-HSS environment. SLF is queried during the Registration and Session Setup to get the address of the HSS maintaining the subscriber specific data.

Inge Grøn­bæk received his M.Sc. degree in computer science from the Norwegian Institute of Technology (NTH/NTNU), Trondheim, Norway in 1977. He has since been involved in protocol design and implementation for systems involving circuit switching, packet switching, and message handling. He was for a period seconded from Siemens Norway A/S to NATO at SHAPE Technical Centre (STC) in the Hague, The Netherlands, where he worked in the area of protocol standardisation. In 1994 he took a position as Chief of technical development at the Networks division of the Norwegian Telecom. In September 1998 he transferred to Telenor (former Norwegian Telecom) R&D as Senior Adviser. In the period 1999 to 2002 he was the Telenor representative in the international 3G.IP Focus Group aiming at applying IP technology in third generation mobile systems. He is currently actively involved in research in the area of IP based network evolution, focusing on routing, mobility management, interoperability and services.

email: inge.gronbak@telenor.com

Voice over WLAN (VoWLAN) – A wireless voice alternative?

TROND ULSETH AND PAAL ENGELSTAD



Trond Ulseth is Senior Research Scientist at Telenor R&D

Voice over WLAN (VoWLAN) is a natural evolution of VoIP. It is also a potential supplement or a potential competitor to 3G mobile systems. This article presents an overview of WLAN technologies and the issues that are relevant for voice communication over WLAN. The basic IEEE 802.11 mechanisms are described, as well as new security and QoS mechanisms recently being standardised by IEEE. Handover issues are addressed. A review of available papers on real-time communication over WLAN highlights the challenges to VoWLAN. There is also a short discussion on 3G-VoWLAN interworking. Finally user equipment aspects and the market potentials are discussed.

Introduction

Wireless Local Area Networks (WLANs) have become very popular in recent years. The most popular technology is based on IEEE Standard 802.11 [1]. 802.11-based WLANs are often considered as a wireless extension to the Ethernet standardised in IEEE Standard 802.3 [2]. WLAN is used both as an alternative and a complement to wired LAN.

With the increasing popularity of IP technology, real-time applications such as interactive two-way voice and multimedia over IP have become popular. A natural evolution of voice over IP is voice over WLAN (VoWLAN).

Like VoIP the growth of the VoWLAN market has been slower than analysts predicted 1-2 years ago. Still it is predicted that the market will take off soon. The main hurdles to VoWLAN seem to be

- Security issues
- WLAN QoS mechanisms
- Handset availability.

New standards on security and WLAN QoS will gradually remove these hurdles. VoWLAN handsets and VoIP software for PDAs with WLAN capability are already available, but few have standardised QoS functionality implemented so far. The WLAN Access Points are potential bottlenecks for VoWLAN networks, and strategies for VoWLAN network management need to be addressed before deploying VoWLAN.

Another issue is the integration of 3GPP systems and WLANs. The data transmission capacity of a WLAN is greater than that of a 3GPP connection. Consequently, WLAN access to IP networks/Internet offers better data throughput than 3GPP access: Where available users prefer WLAN access to 3GPP data access. A feasibility study carried out within the 3GPP project [3] describes six 3GPP-WLAN interworking scenarios.

802.11 standard

Basic mechanisms

IEEE adapted Standard 802.11 [1], which standardised a WLAN technology in 1997. The standard was revised in 1999. Later, several supplements and amendments to the standard have been issued, and there is ongoing work on new supplements and amendments. The standard defines two modes of operation;

- Ad hoc mode, where the stations communicate with each other;
- Infrastructure mode, where an Access Point provides access to another network that may be fixed.

The standard defines the MAC (Medium Access Control) and physical characteristics. At present three physical layer options are standardised:

- IEEE Standard 802.11b [4]. This is the technology that has widest deployment today (end 2005). The theoretical capacity of 802.11b is 11 Mbit/s and the radio frequency band used is 2.4 GHz. The standard specifies the use of Complementary Code Keying (CCK) Modulation.
- A high bit-rate alternative to this standard is specified in IEEE Standard 802.11a [5]. The theoretical capacity of IEEE 802.11a is 54 Mbit/s. The devices operate in the 5 GHz frequency band using Orthogonal Frequency Division Multiplexing (OFDM) technology. The use of the 5 GHz frequency band in Europe caused some problems with regulatory requirements and interference with other services. To overcome this problem, an amended version of IEEE 802.11a, IEEE Standard 802.11h [6] was developed. In Norway (and a number of other European countries), the IEEE 802.11a/h systems can only be used indoors.



Paal E. Engelstad is Research Scientist at Telenor R&D

- A high bitrate alternative based on transmission on the 2.4 GHz radio frequency band is defined in IEEE Standard 802.11g [7]. The OFDM technology is also specified in IEEE 802.11g, and the maximum theoretical bit-rate is 54 Mbit/s. For backward compatibility with 802.11b devices, the standard also specifies support of CCK modulation.

It is important to note that the capacity indicated is the theoretical capacity. In a WLAN network the actual throughput will depend on

- The radio transmission environment;
- The characteristics of the information to be transmitted (i.e. packet size);
- The number of transmitters active.

The actual throughput will therefore be lower than the maximum rate indicated.

802.11b and 802.11g devices can co-exist. Several scenarios are possible:

- *IEEE 802.11g only.* When the Access Point (AP) and all clients support 802.11g, communication occurs at highest possible rate.
- *IEEE 802.11g Access Point, mixed clients.* When the AP supports IEEE 802.11g and there is a mixture of 802.11g and 802.11b clients, the AP instructs the 802.11g clients to use a protection mechanism. The consequence is a reduced throughput, but the throughput for the 802.11g clients is still higher than that for 802.11b clients.
- *IEEE 802.11b Access Point, IEEE 802.11g clients.* The clients will operate in IEEE 802.11b mode.

To manage communication in a mixed b/g environment, protection mechanisms are described in IEEE Standard 802.11g [7]. The principle is to use short messages sent in 802.11b format:

- Request to Send (RTS)/Clear to Send (CTS) messages;
- CTS-to-shelf mechanism where only a Clear to Send message is sent to clear the air.

IEEE has recently initiated work on a new high-speed WLAN standard identified as IEEE Standard

802.11n. It will probably work in the 2.4 GHz frequency band. The transmission capacity is expected to be 2-4 times that of 802.11a/g. An optimistic workplan indicates approval of the standard late 2006.

DCF (Distribution Coordination Function) is the mandatory 802.11 access control function. The DCF uses CSMA/CA (Carrier Sense Multiple Access with Collision Avoidance). Before transmitting, the end-point senses the radio channels to determine if another end-point is transmitting on the same radio channel. The basic DCF signal flow is illustrated in Figure 1. If the radio channel has been sensed idle for a period equal to

$$DIFS + BackoffTimer$$

where

DIFS is defined in Table 2,

$$BackoffTimer = Random() * SlotTime,$$

Random() is a pseudorandom number drawn from a uniform distribution over the interval [0, CW]. CW shall be within the range

$$CW_{min} \leq CW \leq CW_{max}$$

CW_{min} , CW_{max} and *SlotTime* depend on the physical characteristics. They are described in Table 1. The parameters for IEEE 802.11g depend on the working mode, the default *SlotTime* is 20 μ s, but 9 μ s is used when all stations of the BSS¹⁾ support 802.11g.

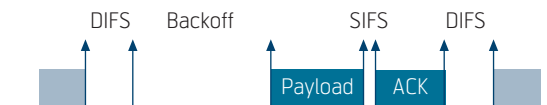


Figure 1 DCF method signal flow

	CW_{min}	CW_{max}	<i>SlotTime</i>
IEEE 802.11a, g	15	1023	9 μ s ²⁾
IEEE 802.11b	31	1023	20 μ s

Table 1 Physical medium dependent parameters of IEEE 802.11 WLAN

1) BSS – Basic Service Set. A set of stations controlled by a single coordination function.

2) For IEEE 802.11g the *SlotTime* is 20 μ s unless all stations in the BSS support the high-speed mode.

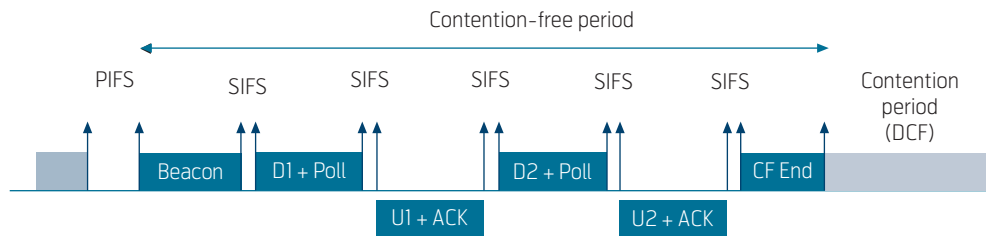


Figure 2 PCF method signal flow

The CW parameter shall initially be chosen in the range $[0, CW_{min}]$. After every unsuccessful transmission attempt the CW parameter is doubled until the value of CW_{max} is exceeded. The CW parameter then remains at this value until it is reset. The CW parameter is reset after each successful transmission.

Each received MSDU (MAC Service Data Unit) requires an ACK message. The time interval between the MSDU and the ACK, and the ACK and the next MSDU shall be equal to SIFS. Two carrier detection mechanisms are identified in [1]:

- Physical mechanism
- Virtual mechanism.

The physical mechanism is described in each relevant supplement [4], [5] and [7]. The virtual mechanism is based on the use of a Duration/ID field to create an internal timer, the Network Allocation Vector (NAV). The Duration/ID field is present in directed frames and in the optional RTS/CTS frames.

When a packet is not received correctly, the MAC layer retransmits the lost packet. Several attempts might be carried out. The default number of retransmissions specified in IEEE Standard 802.11 is 7 for data packets and 4 for RTS packets.

The DCF has no priority mechanism and can be considered as a best effort service. It is found to have

poor performance under heavy and unbalanced load conditions [8].

The PCF (Point Coordination Function) is an optional access method. This method uses a Point Coordinator (PC). The PC shall operate at the Access Point to determine which end-point currently has the right to transmit. The operation is essentially that of polling where the PC is performing the role of polling master. The PCF method signal flow is illustrated in Figure 2.

The PCF provides mechanisms for prioritised access to the wireless medium. Two time intervals, CFP (Contention-Free Period) and CP (Contention Period) are defined. A CFP and the following CP forms a superframe. A superframe must include a CP that is long enough to allow at least one MSDU delivery. During CFP the PCF is used for accessing the medium, while DCF is used during CP. When in CFP mode the CP senses the wireless medium for a period PIFS which is shorter than DIFS but longer than SIFS. If the medium is free, the transmission is initiated. The PC will thus have priority over clients that operate in DCF mode, but does not disturb transmission of ACK frames.

A superframe starts with a beacon management frame transmitted by the AP. Based on this frame all other devices update their NAV. The PC now controls the wireless medium. All clients identified in the CP polling list will then be polled. The PC asks the client to transmit pending frames. Pending frames to the actual client will be included in the polling frame. The client shall acknowledge the successful receipt of the poll and any data. If no response is received within the time interval PIFS (from end of data sent by PC), the next client on the polling list is polled. Therefore no idle period longer than PIFS occurs during PCF.

When the PCF period expires or all the clients on the polling list have been polled, a CF-End frame is sent. The Contention period using DCF as described in the previous section is then initiated. When CP has expired, a new superframe starts. The CP maintains a

DIFS	DCF Interframe Space. $DIFS = SIFS + 2 \times SlotTime$
PIFS	PCF Interframe Space. $PIFS = SIFS + SlotTime$
SIFS	Short Interframe Space. Is 16 μs for IEEE 802.11a and 10 μs for IEEE 802.11b/g
EIFS	Extended Interframe Space. $EIFS = SIFS + t_{ACK} + t_{Preamble} + t_{LCPHeader} + DIFS$
SlotTime	Depends on the physical medium, is 20 μs for IEEE 802.11b and 9 μs for IEEE 802.11a. For IEEE 802.11g the SlotTime is 20 μs except that an optional 9 μs SlotTime may be used when all stations in the BSS consist of high-speed (non 802.11b) stations.

Table 2 Time intervals defined in the IEEE 802.11[1] standard

polling list based on the responses to a contention-free poll.

In the IEEE 802.11 [1] standard a number of time intervals are defined. For clarity these intervals are defined in Table 2.

In addition to the payload there are several headers and control packets that influence the throughput of a WLAN connection. An overview of these headers is presented in Table 3.

For speed adaptation all PHY headers are sent with a rate of 1 Mbit/s when using CCK modulation or 6 Mbit/s when using the OFDM technology. The consequence is that the header related delay is 192 μ s (CCK modulation) or 60 μ s (OFDM).

QoS extensions

The IEEE Standard 802.11e amendment was introduced to support QoS on the 802.11 MAC layer [9]. The basis of 802.11e is the *Hybrid Coordination Function* (HCF), which is the enabler for QoS support. HCF has two medium access mechanisms; *Enhanced Distributed Channel Access* (EDCA) and *HCF Controlled Channel Access* (HCCA). EDCA is used for contention access, while HCCA is for contention-free access controlled by polling from the access point. These relationships are shown in Figure 3.

To understand the principal differences between EDCA and HCCA, one may draw the analogy to IP level DiffServ and IntServ. EDCA, like DiffServ, divides traffic into traffic classes and provides prioritised, qualitative, and relative differentiation between the traffic classes without providing any hard guarantees. On the other hand, HCCA (like IntServ) can handle traffic on a per-application level and provides parameterised, quantitative and “guaranteed” differentiation. Although the “guarantees” of HCCA are harder than those of EDCA, it is always difficult to talk about guarantees when dealing with the inherently unreliable 802.11 medium.

In the following, we will summarize the main characteristics of EDCA and HCCA, and finally turn our attention to other features of 802.11e for QoS support and for improving performance in general.

EDCA

EDCA is an improvement of the DCF medium access mechanism of the base 802.11 specification in Figure 3. It is therefore also referred to as *Enhanced DCF*

Headers and control frames	Size (Bytes)
PHY header	24
MAC header	34
Voice packet headers (RTP/UDP/IP)	40/60 ³⁾
RTS	20
CTS	14
ACK	14
CF Poll	34 ⁴⁾
CF End	20
CF End + ACK	20

Table 3 WLAN headers and control frames

(EDCF). With legacy 802.11 (i.e. without 802.11e), the DCF provides per-station channel access and the post-backoff mechanism of 802.11 ensures fairness between each station. This means that if many stations associate and communicate over an AP, they will all get an approximately equal share of the channel bandwidth. This fairness property is in some ways a very compelling feature of DCF, but not if QoS is to be supported. Another drawback of DCF in terms of QoS is that it specifies the use of only one single transmission queue per station, limiting the opportunities for QoS differentiation even further. With 802.11e, EDCF addresses those two shortcomings of DCF. First, EDCF introduces the possibility of “unfairness”; i.e. that some traffic classes get a higher share of the channel capacity. Second, the traffic classes are split into separate queues.

IEEE Standard 802.1D [10] provides guidance on the mapping of user priorities into *eight* different traffic classes. However, one has decided that differentiation between *four* groups of traffic will be sufficient for 802.11e. 802.11e therefore introduces four classes, also known as Access Categories (ACs), and a trans-

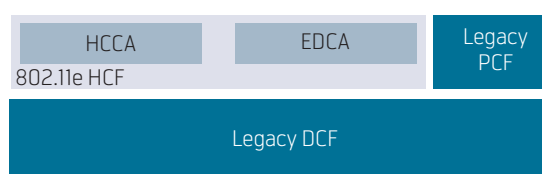


Figure 3 Relation between HCCA and EDCA of 802.11e HCF, and relation to PCF and DCF of legacy 802.11

³⁾ In IPv4 the headers are 40 bytes; in IPv6 the headers are 60 bytes.

⁴⁾ The CF poll frame may or may not contain data.


	(UP - same as 02.1D user priority)	Designation	Category (AC)	(Informative)
lowest  highest	1	BK	AC_BK	Background
	2	-	AC_BK	Background
	0	BE	AC_BE	Best effort
	3	EE	AC_BE	Best effort
	4	CL	AC_VI	Video
	5	VI	AC_VI	Video
	6	VO	AC_VO	Voice
	7	NC	AC_VO	Voice

Figure 4 Recommended mappings of User Priorities to 802.11D traffic classes and to 802.11e Access Categories (ACs)

	AC[3]	AC[2]	AC[1]	AC[0]
AIFSN	2	2	3	7
CW _{min}	3	7	15	15
CW _{max}	15	32	1023	1023
Retry limit (long/short)	7/4	7/4	7/4	7/4

Table 4 Recommended (default) parameter settings for 802.11e when used in combination with 802.11g

mission queue associated with each AC on each station. Figure 4 shows how the aforementioned eight user priorities should be mapped to the four ACs.

EDCF gives different priority to each AC by providing different channel access parameters as shown in Table 4. Since each transmission queue accesses the channel more or less independently from the three other transmission queues and with a different set of access parameters, the differentiation between the ACs are ensured. In Table 4, we see that Voice-over-WLAN traffic is assigned the highest priority.

The traffic class differentiation of EDCF is based on assigning different access parameters to different ACs. First and foremost, a high-priority AC is assigned a minimum contention window that is lower than (or at worst equal to) that of a lower-priority AC. At a lightly loaded (or “unsaturated”) medium, the post-backoff of the high-priority AC will normally be smaller than the post-backoff of a low-priority AC, resulting in an average higher share of the channel capacity. Moreover, as the channel gets more congested (or “saturated”), the high-priority AC will on

average have to refrain from the channel for a shorter period of time than what the low priority AC has to. In other words, while the contention window is fixed without 802.11e, now it is set differently for each AC. Using queue scheduling as an analogy, one may compare differentiation based on contention windows with Weighted Fair Queuing, because of the statistical nature of the weighted backoff process.

Table 4 shows that IEEE recommends a very small minimum contention window for Voice-over-WLAN traffic. The voice traffic gets the highest priority and the delay and jitter due to random access is minimized.

Another important parameter setting is the Arbitrary Interframe Space (AIFS) value, measured as a Short Interframe Space (SIFS) plus an AIFSN number of timeslots. In other words, AIFS is a modification of the DIFS of 802.11. While DIFS uses SIFS plus two additional timeslots, AIFS allows for the use of more than two timeslots. Normally the highest priority AC(s) will be assigned an AIFS that is equal to DIFS, while lower priority ACs are assigned a larger AIFS. In general, a high-priority AC is assigned an AIFSN that is lower than (or at worst equal to) the AIFSN of a lower-priority AC. The most important effect of the AIFSN setting is that the high-priority AC normally will be able to start earlier than a low priority AC to decrement the backoff counter after having been interrupted by a transmission on the channel. At a highly loaded channel where the decrementing of the backoff counter will be interrupted by packet transmissions a large number of times, the backoff countdown of the high-priority AC will occur at a higher average speed than that of the lower-priority AC. As the wireless medium gets more and more congested, the average number of empty timeslots between the frames transmitted by the higher-priority ACs might be lower than the AIFSN value of the low-priority AC. At this point, the AC will not be able to decrement its backoff counter, and all packets will finally be dropped instead of being transmitted. This is referred to as “starvation”. Using queue scheduling as an analogy, one observes that this kind of differentiation can be compared with priority queuing, due to this starvation property.

Table 4 shows that IEEE recommends the smallest possible AIFS value (equal to DIFS) for Voice-over-WLAN. This translates into faster access to the channel due to faster average backoff countdown rate. It also means that the chances of Voice-over-WLAN traffic to experience starvation are minimized.

Other differentiation parameters that may be adjusted in 802.11e (and which are also explicitly or implicitly included in the model proposed below) are the retry

limit (i.e. how many times a packet is attempted to be transmitted before it is dropped), the maximum contention window and the TXOP-limit, namely the maximum length of a “transmission opportunity”.

In 802.11e *Transmission Opportunity (TXOP)* is introduced. While in legacy 802.11 the channel is accessed on a per-frame basis, in 802.11e stations contend for time intervals (TXOPs) where the node is allowed to transmit. The key idea is that short frames, such as Voice-over-WLAN packets, would get an unfairly low share of the channel bandwidth if access were on a per-frame basis. With 802.11e, a QoS Station (QSTA) may for example fill a TXOP with a series of shorter packets, and will not have to contend for the channel for each packet separately.

Another benefit of the TXOP is that QSTAs that are sending on a low bit-rate will not keep the channel busy any longer than stations sending at higher bit rates. The transmission interval that is assigned to a transmission is fixed, and it is up to the QSTA to fill the TXOP with packets or packet fragments.

Although Table 4 listed the recommended parameter values for each AC, the access point (also referred to as “QAP” in 802.11e) may adjust these dynamically. Each periodically transmitted beacon frame may include an Information Element containing these settings. The other stations (also referred to as “QSTAs” in 802.11e) overhear the Beacon and set the access parameters accordingly. The beacon frame also contains information about the number of stations in the BSS, channel utilization and available capacity for the channel and QoS capabilities that are supported at this QAP.

Despite the fact that 802.11e introduces a separate transmission queue per AC, and despite the fact that each AC accesses the channel with different access parameters and with a separate state, it should be noted that access to the channel is not entirely independent. The problem arises when two ACs on a QSTA attempt to access the channel at the same time. 802.11e solves this conflict by allowing the AC of the highest priority to be transmitted while the other AC is going through the backoff procedure as if it experienced a collision on the wireless channel. In other words, there is a separate contention-based mechanism occurring between the transmission queues on each station before packets are transmitted on the wireless medium. This internal conflict resolution mechanism is taken care of by a module called the Virtual Collision Handler as depicted in Figure 5.

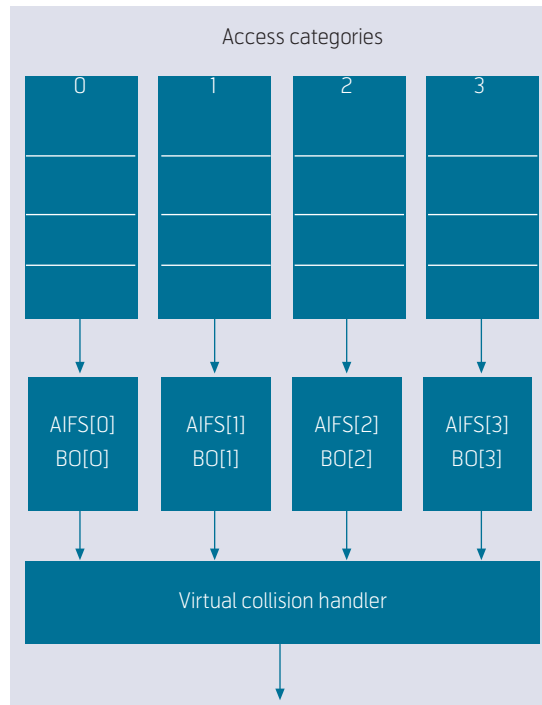


Figure 5 Virtual Collision Handler

HCCA

Just like EDCA is the 802.11e improvement of legacy DCF, HCCA is the 802.11e counterpart of legacy PCF, see Figure 3. Like PCF, HCCA is based on polling. Since an access point (QAP) must be in charge of the polling, HCCA is only possible in an Infrastructure BSS (referred to as QBSS in 802.11e).

HCCA is possible by allocating some parts of the super-frame to polling. In fact, the super-frame concept of 802.11e is the same as that of 802.11, except for the use of TXOPs. The frame starts with a beacon that announces the length of the Contention Period (CP) and Contention-Free Period (CPF). HCCA does not suffer from the same problem as PCF with unpredictable beacon delays, because the QSTAs are not allowed to transmit at the time when the beacon frame should be transmitted. The last frame before the beacon frame must be short enough to finish before the beacon frame.

HCCA is used alone in the CPF but also in CP together with EDCF. Controlled channel access is used in periods called *Controlled Access Periods (CAP)* where polling is done by the HC. *Controlled Contention Intervals (CCI)* may also occur in the CAP where stations only contend for the medium for requesting new TXOPs. The stations can request the TXOPs by sending a special reservation request frame to the HC during this period. By having the highest priority, the HC can allocate TXOP to itself whenever it has a frame to send. The HC waits for the medium to be idle for a PIFS interval of time and

then gives itself a TXOP without entering the backoff procedure. In this way the HC has priority over any other AC or 802.11 station. In the CP period the stations use the EDCF mechanisms to get TXOPs, but can also be polled by the HC where they get a TXOP and can deliver as many MSDUs as fit into to the time allocated by the TXOP. In the CFP period, stations cannot contend for the medium by using EDCF but can only be polled by the HC.

HCCA might be useful for Voice-over-WLAN traffic, especially when the traffic load on the WLAN is high. It requires an appropriate scheduling mechanism implemented for the polling. The algorithm is implementation specific and not covered by the standard.

Other features of 802.11e

Contention Free Bursting and Block ACKs in a TXOP
As mentioned earlier 802.11e allows for sending multiple MSDUs during the TXOP if they fit in the time available. This is called *Contention Free Bursting* (CFB).

Another improvement is that in 802.11e multiple MSDUs can be transmitted in a row without being acknowledged separately. Instead, all the frames are acknowledged with a single ACK. This gives better performance under good channel conditions. The delay from having to acknowledge every MSDU before sending the next is avoided. Block ACKs can be sent either by the end of the TXOP of the transmitted MSDUs or in a later TXOP.

Direct Link Protocol for direct communication between QSTAs

With legacy 802.11 STAs, communication between two STAs belonging to the same BSS relays via the AP. Hence, more than twice as much of the channel capacity is required as compared to direct communication. With the *Direct Link Protocol* (DLP), two 802.11e QSTAs can use the AP to negotiate a direct link between them without having to send traffic via the QAP. However, we believe that the usefulness of this mechanism might be limited for many applications of Voice-over-WLAN traffic. For IP telephony over WLAN, for example, the communicating parties will often be physically located so close together that they might not need to speak on the phone.

Security

The wireless link gives potential attackers an easy access to the transport medium. Security is therefore a challenge to WLANs. WEP (Wired Equivalent Privacy) is a security mechanism specified in IEEE 802.11 [1]. The intention is to provide a security level equivalent to that of a wired LAN. The encryption algorithm used is the RC4 algorithm, a symmetric

algorithm developed in 1987. The standard specifies a 40 bits key. Several suppliers have also implemented a 128 bits key. Although WEP provides security improvements compared with no mechanism, several weaknesses have been discovered. The WEP user authentication solution is also considered insufficient.

IEEE Standard 802.11i [11] has recently been approved by IEEE. In addition to data encryption, the standard provides improved authentication and authorization capabilities. The standard specifies two data encryption mechanisms:

- TKIP (Temporal Key Integrity Protocol). This mechanism is also known as WPA (WiFi Alliance).
- CCMP (Counter mode with CBC-MAC Protocol). This mechanism is also known as WPA2 (WiFi Alliance).

TKIP is considered as a first step towards more robust security solutions. While TKIP uses the RC4 encryption algorithm, CCMP uses the AES (Advanced Encryption Standard) algorithm. AES is a new, strong symmetric encryption algorithm. The key is 128 bits.

There are specified authentication and access control procedures for the WLAN. This mechanism only gives the user access to the authentication server prior to authentication. All other traffic is stopped at the Access Point. The authentication occurs when a client first joins a network. Then, periodically, authentication recurs to verify that the client has not been subverted or spoofed. Although these mechanisms, particularly those specified in IEEE Standard 802.11i [11] (or WPA/WPA2), provide protection to attackers on the radio access, they are not problem free for VoWLAN users. One cryptographic related problem is increased media link delay. Investigations made by Barbieri et al. [12] address these issues. The focus of these investigations is an IPSec implementation. It is however clear that the problems are independent of the security protocols used.

The cryptographic engine may be a bottleneck, particularly when there is a mixture of real-time traffic and non real-time traffic, which usually is the case for Access Points. The design of the QoS and crypto mechanism need to be aligned. Another issue is the increased overhead which requires more processing (and increased delay) in the network elements.

Security issues for VoIP protocols are addressed in [18]. The mechanisms identified may solve the VoWLAN security problems, but the problems

related to the use of the WLAN security mechanisms (i.e. WEP or 802.11i mechanisms) remain unsolved.

Handover and roaming

One of the WLAN mobility related problems is seamless handover and roaming. Before discussing these issues it could be useful to clarify the terminology. The 3rd Generation Partnership Project (3GPP) has published a report where the terminology used within the project is documented [13]. In this report there are two descriptions of *Handover*:

The transfer of a user's connection from one radio channel to another (can be the same or different cell).

Or,

The process in which the radio access network changes the radio transmitters or radio access mode or radio system used to provide the bearer services, while maintaining a defined bearer service QoS.

The description of *Roaming* is,

The ability for a user to function in a serving network different from the home network. The serving network could be a shared network operated by two or more network operator.

These descriptions are not always followed, the term roaming is often used to describe handover. Roaming as described above involves two service providers. The access control mechanisms may be more complicated than those of handover. Roaming issues will not be addressed in this paper.

When a WLAN client is operating in infrastructure mode it will try to associate with an Access Point (AP) in its vicinity. Each AP constitutes a basic service set (BSS) and all the traffic to and from the clients will go via the AP. To cover a larger area, multiple Access Points can be connected via a distribution system (DS) to form an extended service set (ESS). When a client is moving out of the coverage area (cell) of one AP it can reassociate with another AP (within the same ESS). This is handover.

The new Access Point may support different technologies, e.g. Access Point 1 could support 802.11b only while Access Point 2 may support 802.11a only. In this case there is a technology change involved. This handover scenario is often described as *Vertical*

handover, while handover from e.g. one 802.11b Access Point to another 802.11b Access Point often is described as *Horizontal handover*.

Descriptions of the handover process are given in [14] and [15]. During the handover, management frames are exchanged between the mobile station and the AP. A WLAN client cannot associate with two Access Points at the same time. The consequence is that the communication will be interrupted during the handover period, and voice packets are delayed or lost.

A trial-use standard [16] describes recommended practices for implementation of an Inter-Access Point Protocol (IAPP) on a Distribution System (DS). Aironet, Lucent Technologies, and Digital Ocean jointly developed the Inter-Access Point Protocol (IAPP). Among other things, IAPP extends multi-vendor interoperability to the roaming function. It addresses roaming within a single ESS and between two or more ESSs. The 802.11 standard [1] describes two scanning methods to identify APs within the ESS:

- Passive scanning
- Active scanning.

A station using *Passive scanning* switches to the first channel allowed by the regulatory domain and waits for *Beacon*⁵⁾ frames. If the station receives a Beacon frame, it measures the Signal-to-Noise Ratio and stores additional Access Point information. After a specific time, the station switches to the next channel until every channel is scanned. To speed up this process the non-overlapping channels may be scanned. This alternative is called *Fast passive scanning*.

When using *Active scanning* the station attempts to find the network by transmitting probe frames. The station moves to the first channel and waits for the Probe Delay Timer to expire. If an incoming frame is detected, the channel is in use and will be probed. The purpose of this scanning procedure is to find every BSS that the station can join.

Like passive scanning the non-overlapping channels may be scanned to speed up the process (*Fast active scanning*). Simulations of handover between two 802.11b Access Points and using the IAPP protocol have shown that when normal passive scanning is used, the handover delay is more than 600 ms [14]. The normal active scanning procedure is faster, [14] indicates it is in the range between 110 ms and 160 ms depending on the number of stations within the

5) A Beacon frame is a frame the Access Point sends periodically to announce its presence and relay information.

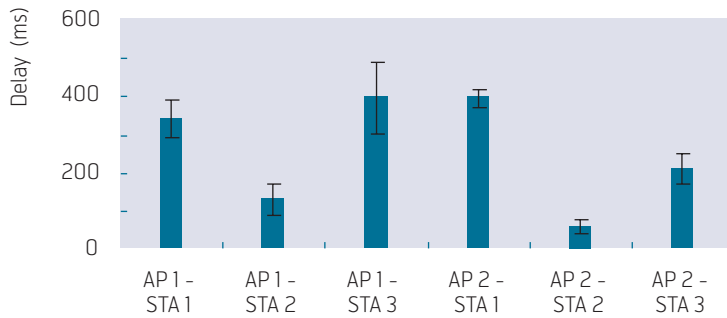


Figure 6 Handover delay reported by Mishra et al. [15]

BSS, while the delay of the fast active scanning procedure may be in the range between 25 and 30 ms.

Practical experiment results reported by Mishra et al. [15] indicate however that these figures may be optimistic. The experiments were carried out on networks using Access Points from two manufacturers using three different PCMCIA cards on the same PC. The handover was repeated several times. Large delay variations were observed. The results and standard deviations are illustrated in Figure 6.

It can be concluded that there are large variations between stations. Even the best combination reported causes loss of two – three 20 ms packets, which is perceptible.

IEEE is currently working on a new standard, IEEE 802.11r. The current workplan indicates publication of the new standard in 2007. One of the goals of the standard is to eliminate perceptible disconnections during handover. Although Roaming is not discussed in this paper, it is worth mentioning that an industry group is working on a roaming protocol, International Roaming Access Protocol (IRAP) [17]. ETSI has decided to use this protocol in the work on standards for Next Generation Networks.

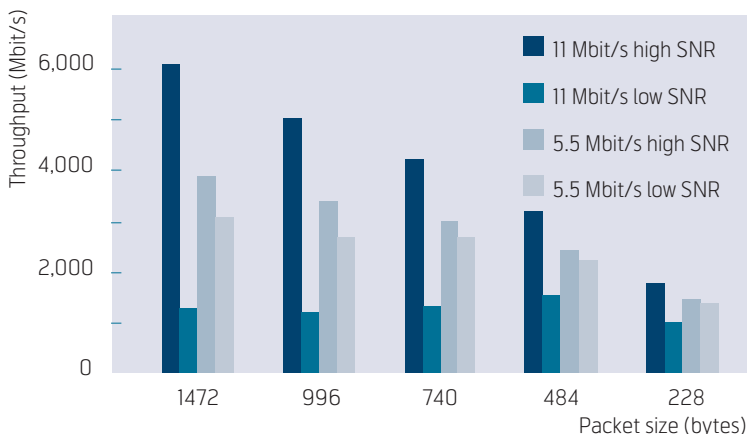


Figure 7 802.11b WLAN throughput measured by Arranz et al.

Real-time communication over WLAN

The basic mechanisms and quality issues for real-time communication over an IP network are discussed in [18] and [19]. Voice (and real-time interactive multimedia) communication is delay sensitive. For that reason voice packets are sent with short intervals, usually 20 ms. The packets are small, between 60 bytes and 200 bytes, depending on the speech coding algorithm chosen, when using IPv4. On the other hand, typical data packets are 1400 – 1500 bytes.

As explained in the section describing the basic IEEE 802.11 mechanisms, there need to be an idle period equal to $DIFS + BackoffTimer$ before a new station can start transmission. The minimum average value of this interval is 360 μ s for 802.11b and 101.5 μ s for 802.11a. In addition to the payload, each packet also contains PHY headers that are transmitted at a low bitrate, the delay introduced by these is 192 μ s when using CCK (802.11b) and 60 μ s when using OFDM.

The transmitting time of a 200 bytes voice packet is approximately 150 μ s when the link rate is 11 Mbit/s and approximately 30 μ s when the link rate is 54 Mbit/s. Compared with the sum of the DCF carrier-sense period and the time used to transmit the PHY headers, the payload transmission may be less than one third of the total time used to transmit a packet. The consequence is that voice communications over WLAN has a large overhead. This is illustrated by Aranz et al. [20] who have measured the throughput of UDP-based transport over an 802.11b WLAN. The measurements are carried out under high communication link (radio) SNR and low (<10 dB) communication link SNR. Other parameters that are varied are the packet size and the WLAN transmission speed. The measurement results for the 11 Mbit/s and 5.5 Mbit/s modes are described in Figure 7. These results illustrate the effect of the overhead, transmission rate and carrier to noise ratio on the throughput. It is also illustrated that lower transmission rates may be beneficial when the radio transmission conditions are unfavourable. However this effect is not large for typical voice packets.

The throughput dependence of packet size means that the number of simultaneous voice connections increases when the packet interval (and the packet size) increases. Figure 8 (802.11b) and Figure 9 (802.11a) based on simulations reported by Garg and Kappes [21] illustrate this effect.

The simulation results also show that for packet intervals of 40 ms or less the choice of speech coding algorithm only has a small effect on the number of voice connections over a WLAN. The figures also

indicate that the maximum number of voice connections over an 802.11a WLAN is about five times higher than the maximum number when transmitting over an 802.11b WLAN. Due to possible co-existence with 802.11b clients the capacity of an 802.11g WLAN is less than an 802.11a WLAN. Results presented by Doufexi et al. [22] show that for 1500 bytes packets and high Carrier to Noise ratio (CNR) in the 54 Mbit/s mode, the throughput of 802.11a is approximately 26 Mbit/s while the throughput when using 802.11g is 17 Mbit/s.

On the other hand, the coverage of 802.11g is better than 802.11a, which means that at the same distance from the Access Point, an 802.11g system may be able to operate in a higher data rate mode than an 802.11a system.

Coupechoux et al. [23] have assessed the maximum number of simultaneous voice calls as a function of the distance between the Access Point (AP) and the client when different speech coding algorithms are used. The WLAN used in the simulations is 802.11b. The results are described in Figure 10. The packet intervals are different for the three codecs; 10 ms for G.711[24], 20 ms for GSM EFR [25] and 30 ms for G.723.1 [26].

It is found that on average, the human voice has a speech activity factor of about 42 %. There are pauses between sentences and words with no speech in either direction. Also voice communication is usually half-duplex; i.e. one person is silent while the other speaks. One can take advantage of these two characteristics to save bandwidth by halting the transmission of cells during these silent periods. This is known as Voice Activity Detection (VAD) or silence suppression.

Implementing VAD/Silence suppression may be beneficial in bottleneck scenarios such as VoWLAN. However, there are problems and limitations. If it is not operating correctly, it can decrease the intelligibility of voice signals and overall conversation quality. When the first utterance is detected after a silence period, it has to switch on transmission quickly in order to avoid loss of information that can make it difficult to understand what is said.

Another important parameter is the hangover time, i.e. the period from silence detection until transmission is stopped. The hangover time should be short in order to improve the efficiency, but it should be long enough to ensure that transmission of low-energy parts of the speech are not stopped. Hums or short utterances like Okay, Yes and so on will reduce the advantage of silence suppression.

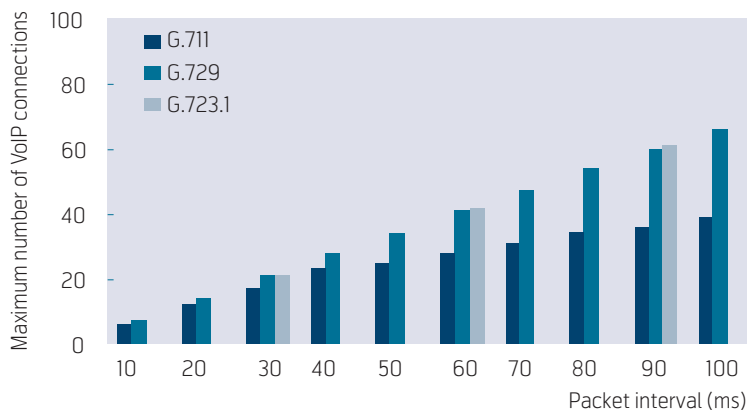


Figure 8 Maximum number of VoIP connections using IEEE 802.11b and different packet intervals

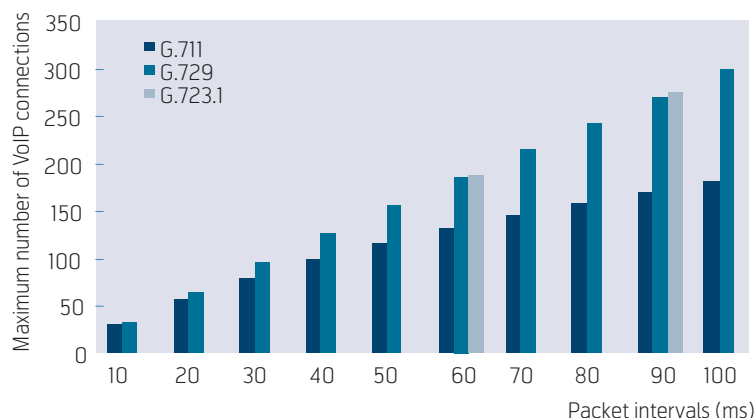


Figure 9 Maximum number of VoIP connections using IEEE 802.11a and different packet intervals

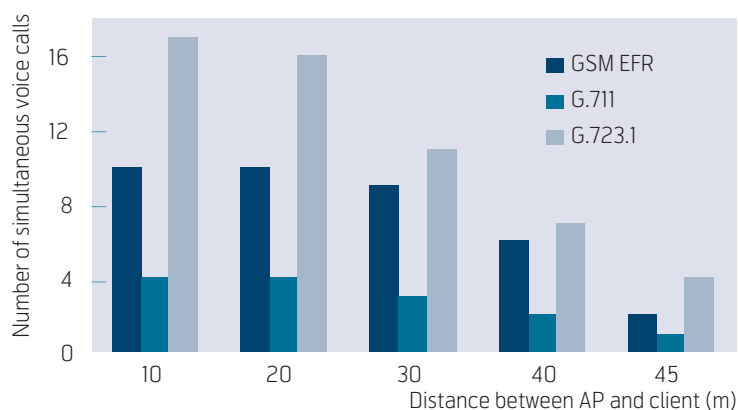


Figure 10 Number of simultaneous connections vs. distance between AP and client, 802.11b

Simulations have been carried out where the average data transmission rate is a variable parameter [27]. The voice codec used was ITU-T Recommendation G.711, the packet intervals were 10 ms, and silence

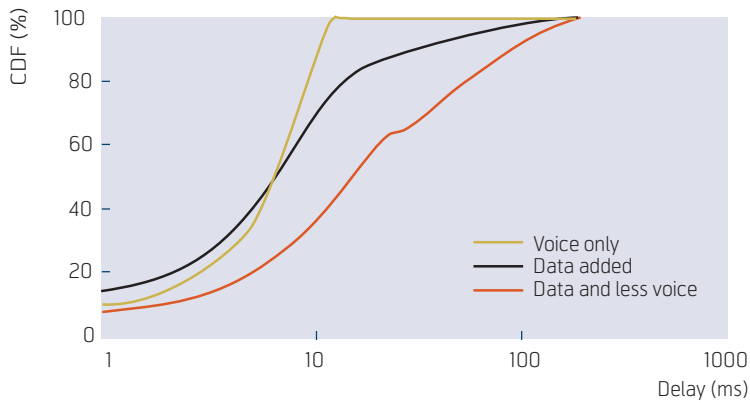


Figure 11 Delay Cumulative Distribution Function IEEE 802.11a

suppression was implemented. The parameter used to define the voice capacity is packet loss equal to or less than 2 %.

Without data traffic an 802.11b WLAN without any voice prioritisation mechanism can support 10 simultaneous voice calls. When 1 Mbit/s additional data traffic is originated in the fixed network and sent to the wireless station the number of simultaneous voice calls is seven. Increasing the data traffic to 2 Mbit/s reduces the number of simultaneous calls to five, while 4 Mbit/s data traffic precludes any voice call due to downlink congestion on the WLAN.

The WLAN related delay for voice packet of 200 bytes or less is small, typically 2 ms, when the traffic is below the maximum capacity of the WLAN.

A Proxim White Paper [28] illustrates the impact of a mixture of voice and data traffic on the access delay. The CDF (Cumulative Distribution Function) of the access delay for three traffic scenarios is illustrated in Figure 11. The scenarios are:

- 1 A wireless channel is fully loaded with voice traffic;
- 2 50 Mbit/s data traffic generated by five additional users;
- 3 Same data traffic as item 2, and the number of voice users is reduced.

The WLAN infrastructure is IEEE 802.11a, and the voice codec bit-rate is 64 kbit/s with 20 ms packet intervals. For a voice only scenario the access delay for 90 % of the packets is less than 10 ms. When data traffic is added, the median delay is about 15 ms, and the maximum delay is approx. 200 ms. When the number of voice users is reduced, the performance in terms of median delay improves, but the maximum delay is still

above 100 ms. Figure 11 also shows that the delay variations (the jitter) increase when data traffic is added.

It is clear that the Access Point is a bottleneck; when there are N stations transmitting the Access Point must access the wireless medium N times for every time each station accesses the wireless medium. The optional PCF mechanism may reduce this problem. Köpsel and Wolisz [29] have evaluated the suitability of the DCF and PCF mechanisms for the transmission of interactive real-time voice. The simulations assume a traffic mix consisting of 15 % real-time traffic and 85 % best-effort traffic. The voice packets are 200 bytes and the packet intervals are 20 ms. The simulations show that the maximum available throughput for DCF mode is about 83 % of the channel bandwidth, while the maximum available throughput for PCF mode is 87–89 % of the channel bandwidth.

The voice capacity of a 2 Mbit/s IEEE 802.11 WLAN is 12 when working in DCF mode and 15 when working in PCF mode. There are a lot of other studies on the use of PCF in a VoWLAN scenario. However, although some of these indicate that there might be a capacity improvement, almost no commercial WLAN product has this option implemented. The use of PCF is therefore more an academic issue than a real life issue.

3G voice and VoWLAN – Competitors or partners?

The data communication capacity of the 3rd Generation mobile system is limited. Where available WLAN access may improve the data communication capabilities. The complementary use of WLAN is considered as an extension of 3G data services.

A feasibility study carried out within the 3GPP project [3] describes six 3GPP-WLAN interworking scenarios. Table 5 presents these scenarios and related capabilities. The scope of this study is not limited to data communication, and includes the whole range of 3G services.

The scenarios most relevant to VoWLAN are the packet-based real-time services of 3G, i.e. scenarios 3, 4 and 5 of Table 5. The scenario extends the 3G IMS services to the WLAN; it is however a matter of implementation whether all services or a subset of the services are provided. Scenario 3 does not provide service continuity. When a user changes access (e.g. from 3G to WLAN), the sessions need to be re-established. Scenario 4 allows the services to survive a change of access between WLAN and 3G systems, but the change of access may be noticeable to the user. In the fifth scenario aspects such as data loss and break time during the

Scenarios	Scenario 1: Common billing and customer care	Scenario 2: 3GPP system based access control and charging	Scenario 3: Access to 3GPP system PS based services	Scenario 4: Service continuity	Scenario 5: Seamless services	Scenario 6 Access to 3GPP system CS based services
Common billing	X	X	X	X	X	X
Common customer care	X	X	X	X	X	X
3GPP system based Access Control		X	X	X	X	X
3GPP system based Access Charging		X	X	X	X	X
Access to 3GPP system PS based services from WLAN			X	X	X	X
Service Continuity				X	X	X
Seamless Service Continuity					X	X
Access to 3GPP system CS based Services with seamless mobility						X

Table 5 3G-WLAN Scenarios and capabilities identified by 3GPP [3]

switch between access technologies are similar to the interruption perceived during an internal 3G handover.

However, these functionalities will hardly be implemented on a short term. Compared with the increased data communication capacity VoWLAN-3G interworking does not provide any technical advantage to the user, the only possible advantage could be economic (i.e. cheaper connection).

User equipment

An increasing number of laptop PCs and PDAs now have WLAN support in its standard configuration. Like VoIP over fixed IP networks solutions for VoWLAN applications can be based on software client implemented on these devices.

There are however limitations related to the use of PCs as voice clients. The reliability of a PC and its software has not yet reached that of a telephone. A strong argument for using VoWLAN is mobility. A laptop PC is mobile, but it is not very convenient to use it as a mobile telephone device.

The PDAs may be a better alternative, but tests carried out indicate that some PDAs may have performance problems and do not fully meet the real-time requirements of voice communication. The problem may not necessarily be the PDA itself, but the Operating System.

WLAN telephones have been on the market for more than five years. The most successful supplier seems to be Spectralink, which had 70 % of the market in 2002.

The company has established partnership with several other suppliers such as Avaya. It is important to note that Spectralink offers a proprietary QoS solution. The solution requires that Spectralink supplies both the handset and the Access Point. One market segment they address is Healthcare. The Spectralink products are ITU-T Recommendation H.323 compliant.

Another H.323-based WLAN telephone that has been on the market for quite a while is Symbol NetVision Phone. However, Symbol has recently announced that this product is discontinued. This is probably linked to an introduction of a new family of devices called Enterprise Digital Assistant (EDA). These devices have WLAN access and integrated VoIP functionality.

There are now companies in the Far East offering 801.11b-based handsets at less than 200 \$. Most VoWLAN handsets support WEP security, and a few handsets support WPA/WPA2 security. As already indicated Spectralink offers a proprietary QoS solution. There are also handset or PDA soft clients that claim conformance with the WiFi Alliance WMM standard, the forerunner of IEEE 802.11e. Some suppliers are also claiming that their product is prepared for support of QoS by a software upgrade.

It is expected that the number of VoWLAN handsets will increase rapidly. Broadcom has introduced a chipset solution for the development of a wireless VoIP handset that will support voice and data applications (web browsing, email, and instant messaging) within an 802.11g WLAN environment. The chips include integrated support for advanced security and QoS.

Texas delivers a WLAN IP Phone Solution and describes a reference implementation. It is believed that this would contribute to cheaper VoWLAN handsets. This trend is strengthened by the TIA/EIA decision to develop a standard for VoWLAN [30].

It is also predicted that future mobile devices will support multiple wireless standards [31]. An obvious combination is 3G and WLAN. The main benefit of a 3G/WLAN device is the increased data communication capacity when accessing a WLAN. The benefits for voice are less obvious, but use scenarios where a voice call is made simultaneously with data communication is an example where the users may benefit from VoWLAN. Market analysts have indicated that by 2009, roughly 30 % worldwide of all cellular voice devices will have some type of WLAN, and roughly 18 % of cellular voice devices will incorporate WLAN designed to carry voice [32].

The future – Will VoWLAN be a success (factor)?

Voice over WLAN is based on two independent technologies:

- Wireless LAN
- Voice over IP.

The success of these two technologies is the first condition for the success of VoWLAN. In spite of security shortcomings, WLAN has already been a success, both in the enterprise market and the residential market. There is an increasing number of VoIP service providers. In Norway most of the ISPs offer a VoIP service. There are also companies such as Vonage in the US and Telio in Norway that provide a VoIP service independent of the ISP. Although the number of VoIP customers is small (a few per cent) compared to the number of PSTN customers, it is increasing. VoIP can hardly be declared a success at the time being (end 2005), but there are strong indications that the future is bright.

The conditions above have to be met but are not sufficient for the success of VoWLAN. Other important elements are

- WLAN QoS availability and the adaptation of the WLAN technology to voice applications;
- Available security and access control mechanisms;
- Available handsets at an acceptable price;
- Usability aspects.

The new QoS and security standards discussed in this paper are well under way. The Wi-Fi Alliance has introduced certification programs for both QoS

(WMM) and security (WPA/WPA2). However, although the standards provide the mechanisms for good implementations, there are several aspects to be considered before a product can be considered to fully meet the user requirements.

Available handsets at an acceptable price are expected in the near future; chip manufacturers are introducing VoWLAN chipsets that include integrated support for advanced security and QoS. This will help reduce the cost of VoWLAN handsets. However, although prices are declining, analysts are not expecting prices to dip below \$300 until 2007 in the US market. As a matter of fact there are indications that the analysts have been too pessimistic when making this estimate, handsets costing less than \$200 are available today, albeit without QoS support and with security mechanisms that are not as specified in IEEE 802.11i.

The usability is linked to security solutions and how an acceptable QoS is obtained. Mobility is assessed as one of the key drivers. By providing VoWLAN technology to health-care workers, for example, hospitals can improve the productivity of doctors and nurses, ensuring they spend time with patients instead of running to retrieve messages. Other market segments that have been highlighted are education, manufacturing plants and retail stores. A common aspect of these market segments is the mobility of the staff. Analysts predict an exponential growth of VoWLAN devices the next four years. A factor that may influence the growth is the services offered by mobile operators to enterprises. In Norway the mobile operators offer a subscription scheme where the communication between the employees is not charged. The scheme weakens the arguments for a VoWLAN enterprise network.

Although most of the analyses have focused on the business market and large companies in particular, there are indications that the growth of the residential market could be larger. One potential obstacle to the VoWLAN success in the enterprise and hotspot markets is the limited capacity of the 2.4 GHz frequency band. In this band there are only three non-overlapping channels. Using 802.11g or the coming 802.11n increases the capacity of the network, but for enterprise networks this is a potential problem. Using the 5 GHz band may solve this problem, but there are restrictions on open air use, and the area that is covered by a 5 GHz Access Point is less than the coverage area of a 2.4 GHz Access Point. This is probably a small problem for the residential market.

As prices decrease, it is predicted that more dual-mode WLAN/cellular handsets will reach the market, enabling enterprise users and consumers to roam

across wireless home networks, the corporate wireless LAN, public WLAN hotspots and mobile networks. This may bring both opportunities and threats for the mobile operators.

To conclude, VoWLAN as a supplement to the fixed network VoIP installation will be a success in the residential market. There are more question marks about the enterprise market where there may be both technical and commercial obstacles to the success of VoWLAN.

References

- 1 IEEE. *Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks. Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications*. Piscataway, NJ, 1999. (IEEE Standard 802.11)
- 2 IEEE. *Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks. Specific requirements. Part 3: Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specification*. Piscataway, NJ, 1999. (IEEE Standard 802.3)
- 3 ETSI. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspect; Feasibility Study on 3GPP system to Wireless Local Area Network (WLAN) interworking (Release 6)*. Sophia Antipolis, 2003. (3GPP TR 22.934 v6.2.0)
- 4 IEEE. *Supplement to Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks. Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. High-Speed Physical Layer Extension in the 2.4 GHz Band*. Piscataway, NJ, 1999. (IEEE Standard 802.11b)
- 5 IEEE. *Supplement to Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks. Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. Higher-Speed Physical Layer in the 5 GHz Band*. Piscataway, NJ, 1999. (IEEE Standard 802.11a)
- 6 IEEE. *IEEE Standard for Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks. Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. Amendment 5: Spectrum and Transmit Power management in the 5 GHz Band in Europe*. Piscataway, NJ, 2003. (IEEE Standard 802.11h)
- 7 IEEE. *Supplement to Information technology – Telecommunications and information exchange between systems – Local and metropolitan area networks. Specific requirements. Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. Amendment 4: Further Higher Data Rate Extension in the 2.4 GHz Band*. Piscataway, NJ, 2003. (IEEE Standard 802.11g)
- 8 Bianchi, G. Performance analysis of IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communication*, 18 (3), 535–547, 2003.
- 9 IEEE. *Amendment to Standard for Information Technology – Telecommunications and Information Exchange Between Systems – LAN/MAN Specific Requirements – Part 11 Wireless Medium Access Control (MAC) and Physical Layer (PHY) specifications: Amendment 7: Medium Access Control (MAC) Quality of Service (QoS) Enhancements*. Piscataway, NJ, October 2005. (IEEE Standard 802.11e)
- 10 IEEE. *IEEE standard for Local and metropolitan area networks Media Access Control (MAC) Bridges*. Piscataway, NJ, June 2004. (IEEE Standard 802.1D)
- 11 IEEE. *Amendment to IEEE Std 802.11, 1999 Edition (Reaff 2003). IEEE Standard for Information technology – Telecommunications and information exchange between system – Local and metropolitan area networks. Specific requirements – Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications – Amendment 6: Medium Access Control (MAC) Security Enhancements*. Piscataway, NJ, 2004. (IEEE Standard 802.11i)
- 12 Barbieri, R, Bruschi, D, Rostu, R. Voice over IPsec: Analysis and Solutions. *18th Annual Computer Security Applications Conference (ACSAC'02)*, Las Vegas, 29–31 July 2002.
- 13 ETSI. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspect; Vocabulary for 3GPP Specifications*. Sophia Antipolis, 2005. (3GPP TR 21.905 v6.9.0)

- 14 Pries, R, Heck, K. *Performance Comparison of Handover Mechanisms in Wireless LAN Networks*. University of Würzburg, Department of Computer Science, 2004. (Report 339)
- 15 Mishra, A, Shin, M, Arbaugh, M. An Empirical Analysis of the IEEE 802.11 MAC Layer Hand-off Process. *ACM SIGCOMM Computer Communication Review*, 33 (2), 2003.
- 16 IEEE. *IEEE Trial-Use Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11™ Operation*. Piscataway, NJ, 2003. (IEEE Standard 802.11f)
- 17 *International Roaming Access Protocols (IRAP) Program*. 2005, January 26 [online] URL: <http://www.goirap.org/>
- 18 Ulseth, T, Stafnes, F. Real-time communication on IP networks. *Teletronikk*, 102 (1), 3–22, 2006. (This issue)
- 19 Ulseth, T, Stafnes, F. VoIP speech quality – Better than PSTN? *Teletronikk*, 102 (1), 119–129, 2006. (This issue)
- 20 Arranz, M G, Agüero R, Muñoz, L, Mähönen, P. Behaviour of UDP-Based Applications over IEEE 802.11 Wireless Networks. *IEEE 12th International Symposium on Personal, Indoor and Mobile communication*, San Diego, 30 September – 3 October 2001.
- 21 Garg, S, Kappes, M. Can I add a VoIP call? *IEEE International Conference on Communications 2003 (ICC'03)*. Anchorage, Alaska, 11–15 May 2003. (Also available as Avaya Technical Report ALR-2002-012, however with a different title.)
- 22 Doufexi, A, Armour, S, Lee, B-S, Nix, A, Bull, D. An Evaluation of the Performance of IEEE 802.11a and 802.11g Wireless Local Area Networks in a Corporate Office Environment. *IEEE International Conference on Communications 2003 (ICC'03)*, Anchorage, Alaska, 11–15 May 2003.
- 23 Coupechoux, M, Kumar, V, Brignol, L. Voice over IEEE 802.11b capacity. *16th ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Networks*, Antwerp, Belgium, 31 August – 2 September 2004.
- 24 ITU-T. *Pulse Code Modulation (PCM) of voice frequencies*. Geneva, 1988. (ITU-T Recommendation G.711)
- 25 ETSI. *Digital cellular telecommunications system (Phase 2+); Enhanced full rate (EFR) speech transcoding; (GSM 06.60 version 8.0.1 Release 1999)*. Sophia Antipolis, 2000. (ETSI EN 300 726 version 8.0.1 Release 99)
- 26 ITU-T. *Dual rate speech coder for multimedia communication transmitting at 5.3 and 6.3 kbit/s*. Geneva, 1998. (ITU-T Recommendation G.723.1)
- 27 Anjum, F et al. Voice Performance in WLAN Networks – An Experimental Study. *GLOBE-COM 2003*, San Francisco, California, 1–5 December 2003.
- 28 Proxim. *White Paper: Voice-Over Wi-Fi Capacity Planning* (2005, July 5) [online] <http://www.proxim.com/learn/library/whitepapers/>
- 29 Köpsel, A, Wolisz, A. Voice transmission in an IEEE 802.11 WLAN based access network. *Fourth International Workshop on Wireless Mobile Multimedia 2001 (WoWMoM 2001)*, Rome, 21 July 2001.
- 30 *Telecommunications Industry Association: TIA to Create New Voice Transmission Standard for Wireless Lan (WLAN) Internet Telephony*. (2004, October 22) [online] URL: http://www.tiaonline.org/media/press_releases/index.cfm?parelease=03-51.
- 31 Chevillat, P, Schott, W. The role of radio LANs in the wireless evolution. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2002)*, Lisbon, 15–18 September 2002.
- 32 *3G* (2004, October 29) [online] URL: <http://www.3g.co.uk/PR/June2004/7864.htm>.

For a presentation of Trond Ulseth, please turn to page 2.

For a presentation of Paal Engelstad, please turn to page 64.

MMoIP – Quality of service in multi-provider settings

TERJE JENSEN



Terje Jensen is Senior Research Scientist at Telenor Research and Development

Considering the discussion and interest on IP Quality of Service (QoS), one might expect that adequate solutions are up and running commercially. The fact is, however, that no full-blown commercial implementations have been introduced regarding ensured and differentiated IP service levels end-to-end. So, why is that? What still needs to be developed? And, what may one do with the results available? This article addresses some challenges, service levels, mechanisms and potential gains related to proper managing QoS for multimedia-over-IP services.

1 Introduction

The Internet is commonly described with great enthusiasm, pointing to the high growth in number of users/servers connected and the traffic carried. Moreover, one can recognize that several of the solutions and protocols developed by one of the central bodies, the Internet Engineering Task Force (IETF), have been adopted in system architectures developed by other bodies. This may also be seen as an acknowledgement of these solutions. From this follows that IP-based solutions would find their positions widened; emerging in systems having little to do with IP and Internet so far.

A fundamental requirement for this to take place, however, is that the *users are pleased with the functionality and service levels presented* – for the wide range of services and corresponding applications that appear. QoS-related mechanisms are aiming for just this: *to maintain the service levels in accordance with the levels that are to be expected*. Examples of service level parameters are session set-up time, availability, information transfer time, and so forth. Importance of the different parameters varies for the different services and how these services are used.

Returning to the genuine Internet design, so-called single-class best-effort service was the sole model present. Although being a simple model – all user requests are treated in the same manner – it is not able to match the wide spectre of applications seen. As long as most of the private users applied e-mail, browsing and file-sharing, there were commonly low requirements on performance parameters beyond throughput; perhaps except dependability parameters. With the emergence of voice and video, and other so-called real-time services, the importance of other parameters grows. This strengthens requirements on QoS mechanisms to ensure that end-users see the proper service levels.

An additional challenge for the commercial service offering is that several providers are commonly

involved. One example is a user starting a Video-on-Demand (VoD) service; where different companies deliver i) the access network, ii) the core/transport network, iii) the video server part, and, iv) the payment and customer aspects. Another example is shown in Figure 1. Here a user may, with the same handset, be utilising different access networks – say private WLAN over DSL, 3G and public WLAN during the same session with a voice call being handed over between these access networks. IP would be a common bearer protocol for this configuration to work smoothly. These configurations imply that the QoS has to be maintained across the corresponding domains or systems. Hence, mechanisms must be devised to take care of the multi-provider environments.

IP capabilities are introduced in steadily more devices, for example, power utility reader, TV sets, remote loggers. This adds to the range of applications and the traffic characteristics meeting the network. The convergence trend is also strong these days allowing for communicating between these devices, and providing consistent user interfaces across different access types, see Figure 1. The fairly low cost of introducing WLAN into a device means that such access points can be attached to different public networks enabling seamless user experiences.

Chapter 2 gives an overall discussion of QoS principles. Multimedia services are briefly presented in the subsequent chapter with the main emphasis on traffic and QoS characteristics. QoS mechanisms are presented in Chapter 4, prior to outlining multi-provider aspects in Chapter 5.

2 QoS – never ending story?

2.1 Overall

In order to provide and configure the network resources it is vital for a network operator to assess characteristics of services to be provided. This also

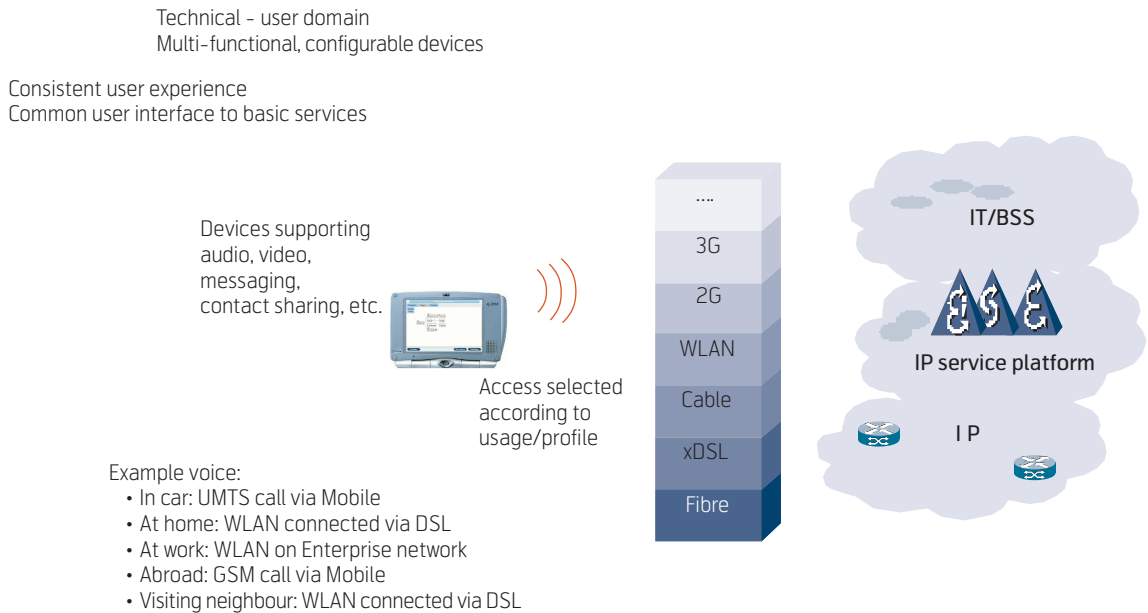


Figure 1 Using the same handset to access services through different access networks – striving for seamless and consistent user experiences

includes QoS requirements of services. Besides used as input when designing systems, guidelines on conditions to place in Service Level Agreements are obtained.

Figure 2 shows four technical areas related to QoS mechanisms:

- Service Level Agreements; addressing relations between two entities in the service provision chain;
- Admission control/conditioning commonly applied at the border of a domain to answer whether more traffic can be admitted and verify that the admitted traffic confirms to announced characteristics;
- Scheduling and resource management configuring the different resource types and handling the packets/traffic classes;

- Signalling/service interfaces between entities to inform the other on change of traffic conditions or QoS requirements.

All these types have to be present in a commercially operating network to be able to ensure service levels. Note that aspects of monitoring are attached to all of these.

QoS should be discussed in view of every facet of service provision, including the relations between a user and a provider. In fact this is where several of the existing IP QoS-related discussions are somewhat narrow – not including all essential phases of the service provision. Besides the strictly technical aspects related to IP packet flows, the more human-related aspects would in most cases be even more important.

Hence, in addition to examining the IP packet transport service more user-centred services must be considered, such as a videoconference, multimedia application sharing, etc. Moreover, the overall customer relation must also be taken into account.

2.2 QoS – definition

As a basic approach to the QoS topic some fundamental results have been elaborated jointly with other European operators (ref. EURESCOM P806, [EU.P806]). The results have been published at different conferences and also provided one of the main fundaments for ITU-T Recommendation E.860, [E.860]. The scope and motivation for that work was to solve the “generic QoS understanding” in a multi-provider environment also considering the multi-ser-

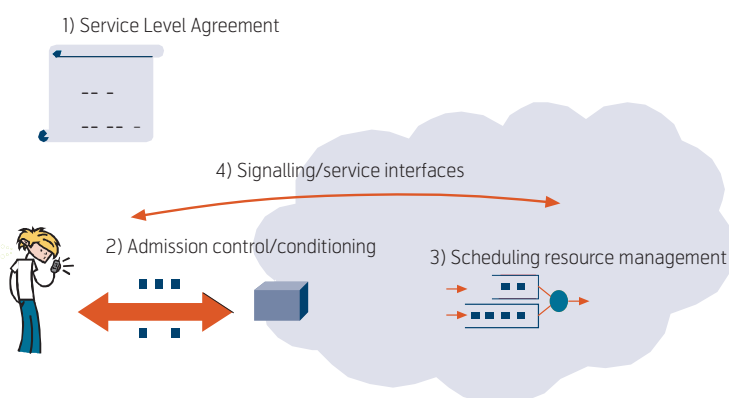


Figure 2 Four technical element categories for QoS

vice and multi-technology setting. Hence, a rather fundamental and generalized interpretation of QoS was needed. In fact the definition chosen – *QoS = degree of conformance of the service delivered to a user by a provider, with an agreement between them* – brings the quality understanding and management for internet/telecommunication in alignment with other industries. It also straightens the confusion between service levels / service classes and QoS. Working in a commercial environment it is important to arrive at a clear interpretation of such essential terms as QoS.

Another important element in describing service characteristics is defining components of services. At a higher level, a service that a user is facing would likely consist of a number of components – each with its specific characteristics. Addressing this area in an efficient manner, a framework for composing services is asked for. The full-blown provider situation has to cover all aspects from advertising and marketing to operation and customer complaint handling. However, one mostly focuses on the network- and operational-related aspects.

One example of composition is a multimedia session that could well be composed of a video component, an audio component and a number of data components. Again, each of these components could have different characteristics. An end-user would frequently relate to the composite behaviour of the components, which make up the complete service.

2.3 Brief on parameter types and mechanisms

Some support for estimating service characteristics is found in publications, including standardisation documents, e.g. 3GPP and ITU. A main challenge seems to be not finding relevant material, but rather to present the requirements in a systematic manner. Typical QoS requirements can be divided into:

- delay-related
- loss-related
- dependability-related.

All these have to be considered, although the third area is less frequently covered in standardisation documents. One example is ITU-T Recommendation Y.1541, ref. [Y.1541] where the following parameters are specified:

- IP packet transfer delay
- IP packet delay variation
- IP packet loss ratio
- IP packet error ratio.

Parameterisation of service components is then possible, both considering usage situations as well as how the services are implemented. In some cases no strict bounds are given for services, hence allowing flexibility in the service delivery. An example is the throughput provided for a TCP session. For dimensioning purposes the application/usage of such services must be considered, that is taking into account that some minimum service levels are commonly expected.

Taking on the user perspective, there is less interest for details of implementation. Assessing proper performance levels asks for further insight into how users experience the services. According to [G.1010] the parameters should comply with:

- Taking into account all aspects of the service from the users' point of view;
- Focusing on user-perceivable effects, rather than their causes within a network;
- Independent of specific network architecture or technology;
- Objectively or subjectively measured at the service access point;
- Easily related to network performance parameters;
- Assured to a user by the service provider, e.g. through Service Level Agreement.

Note that in a multi-provider environment, a given provider may be a user of services delivered by another provider.

In order to ensure IP service levels, the general groups shown in Figure 2 have to be further defined and allocated to different elements as illustrated in Figure 3.

2.4 QoS; why – or why not?

In view of the great commercial interests in IP services, it is natural to raise the question of why the QoS aspects have not been fully deployed already. Some of the arguments for this are:

- The QoS mechanisms are too complex: as described above, functionality is needed in user devices, applications, network elements, support systems, and so forth. This can be considered as a major step, challenging the technical capabilities as well as operational competence. Performance of equipment may also be severely degraded when the mechanisms are activated.

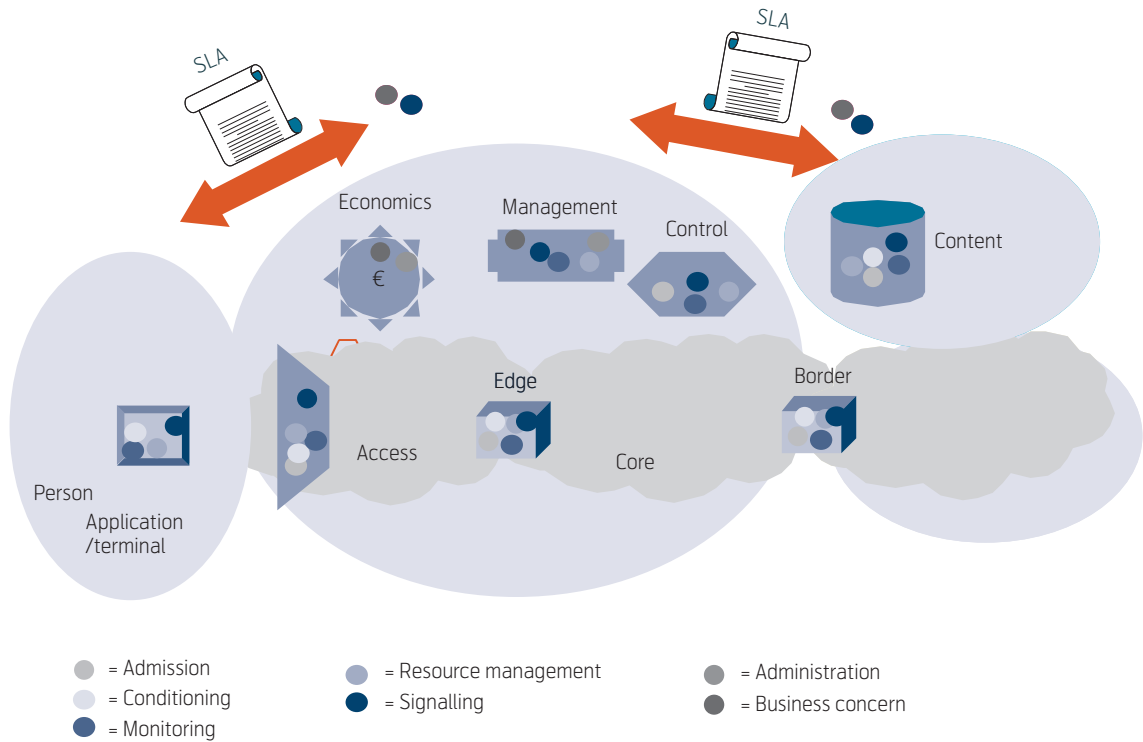


Figure 3 Illustration of IP QoS mechanisms in different elements

- Too costly: looking at equipment prices the differences between supporting a single best-effort class and supporting full-fledged QoS mechanisms are still quite big. Besides the equipment costs, operational procedures and insight into how to tune the parameters have to be taken into account.
- No actual need for ensured service levels: the Internet has been running quite well without the mechanisms described above.
- Simply adding more capacity (“throwing bandwidth at the problem”): in case a performance bottleneck appears, the capacity can be increased until the degradation issues dissolve.
- Similar services to all, mimicking some social fairness principles.

The presence of multimedia services as well as increased commercial pressure on the IP-based services, however, go against these factors:

- Customer and application plurality manifested. A simple technical example is the presence of two transport protocols; TCP and UDP. Although both are carried by IP they imply quite different behaviour on the traffic characteristics.
- IP transport capability became business mission critical. Imagine a travel agency or broker basing

almost all its operation on IP-based services (voice and Web), being out of service for some time would likely also imply being out of business.

- Commercial interests entered (how to increase your margins?): how to provide ensured end-to-end service levels with cost and quality which customers are prepared to buy?
- Equipment matured and insight gained on how to utilise the mechanisms and tune parameters in order to achieve the service operation levels. Cost-wise such solutions compete well with the adding-more-bandwidth-approach for several of the bottleneck cases as shown by a few examples later.

Devising an evaluation of QoS mechanisms the overall situation should be looked into. That is, information transfer, control/management, business issues must be detailed – assessing both “benefits” and “expenses”.

3 Multimedia-over-IP

3.1 Overall

As shown in Figure 4 quite a few elements may be involved in a multimedia session. Two central aspects are the user data transfer and the service control. Different implementations of these exist; two main groups characterised by whether the network is

involved in service control. For several applications only the end-points/-terminals are aware of the multimedia session, leaving the network out of the control loop. A result is that service differentiation and ensured services may be challenging. However, this is a very common situation so far. There is quite a lot of interest on the providers' side to get involved in the service control allowing for a wider range of service offerings. This also implies that service control mechanisms must be introduced, where initiatives such as ETSI-TISPAN come in handy ([Ross06]).

There are several ways of grouping media streams, one being volume-strict versus time-strict. The former refers to media components where the volume is fixed (e.g. file transfers), while the latter refers to sessions having clear start and end instances where volume to transfer is more flexible (such as voice conversations). These also have different requirements on QoS parameters; commonly the former having stricter loss requirements while the latter having stricter delay requirements. Dependability requirements are present for all types.

3.2 Source models and characteristics

Models and characteristics should be devised for all traffic flows in a multimedia session. Illustrations of source models for voice, video and data are depicted in Figure 5. The IP packet arrivals from a source depend on the source behaviour and the coding/handling in the end-systems.

For example for voice, activation of silence suppression can avoid transmission during pauses. In this way, an on-off traffic pattern is generated, where the peak rate remains unchanged but the mean rate is smaller. Typically the mean rate is reduced to less than 50 %. Rather than for single sources, the reduction of the mean bit rate by silence suppression can then be exploited for an aggregation of many voice sources due to the statistical multiplexing effect.

A codec is the unit transforming between analogue and digital format. Several codecs have been defined for voice [Ulse06]. Besides resulting in different bandwidth requirements on the IP layer, they have also different packet characteristics in terms of packet lengths and arrival processes.

Commonly, there is a *trade-off between low bit rate and low delay of a codec*. This comes from overhead added from protocol headers. Although all coding delays are in the tolerable range for the conversational service class, this is only a single delay component, which adds to propagation and queuing delays.

Similarly as for voice, video codecs decide how the frames are generated. For some codecs frames of video arrive at regular intervals determined by the number of frames per second. Each frame is decomposed into a fixed number of slices, each transmitted as a single packet. Encoding delay at the video encoder introduces delay intervals between the packets of a frame.

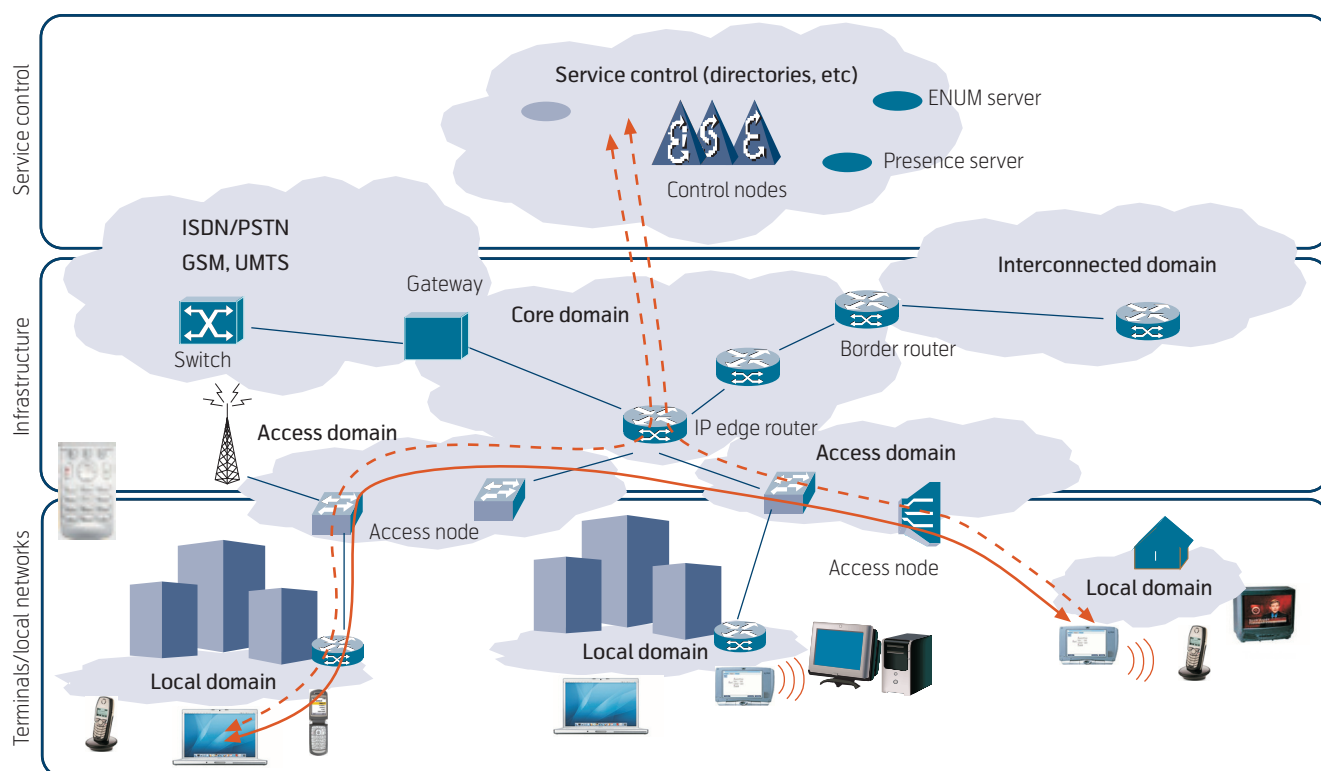


Figure 4 Illustration of elements involved in multimedia-over-IP session

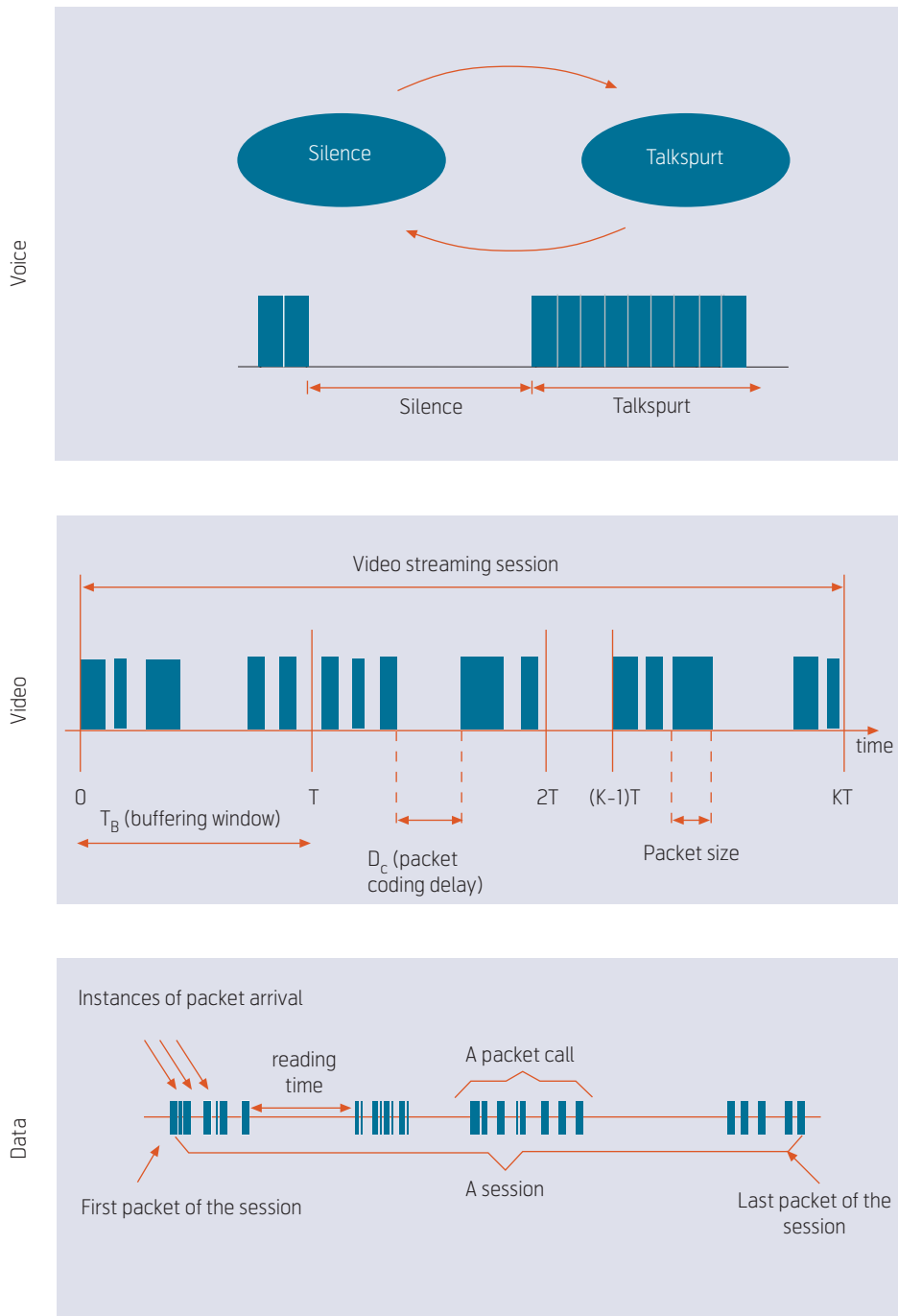


Figure 5 Illustration of source models

A de-jitter buffer window at the receiver is used to restore a continuous display of video streaming data.

Similarities between voice and video characterisation are evident, with packet generation processes heavily influenced by codec engaged. However, different advance video codec schemes introduce more differences. MPEG coding of video causes traffic variability on several time scales including periodically changing rates in predefined groups of pictures (GoP) with frames at different compression levels, where each GoP has the length of usually less than a second, then:

- Variable intensity of motion within a scene impacts the redundancy versus coding gain in subsequent pictures causing short range dependency with correlation in the time scale of seconds up to minutes;
- Variability of the compression gain in different scenes on a larger time scale introducing long range dependency of the traffic rates.

Figure 6 shows variability of real time video traffic for Jurassic Park for 7500 GoPs of a 1.5 hour trace. Due to coding and different efficiency of compression depending on the motion in scenes, there is no

unique value or range for the parameters of video traffic. Video source rates may vary from several Gb/s in HDTV down to 64 kbit/s with a high compression ratio.

For both voice and video, one would define different parameter values for different service types. Examples of service types are:

- Conversation, between a number of parties in “real time”;
- Messaging, referring to a sender initiating an audio/video message to be (eventually) delivered to a set of receivers; and
- Streaming, referring to downloading an audio/video sequence.

There is a wide range of data applications, some describing a human user while others have a machine as trigger. Browsing is an example of the former where a user “clicks-and-downloads” a number of objects due to activated links. Transferring e-mails between servers is an example of the latter. Other data applications send still images, interactive games, instant messaging, transfer bulk data, and payment/point-of-sale. Naturally, these do have different source characteristics. Key parameters for data applications are size of files to be transferred, initiation rate of sessions, and think time between transfers within a session.

When human users are behind the application, on-off traffic is commonly observed. Examples are web browsing and other interactive applications with data transfer for requests and response in opposite directions.

3.3 QoS requirements

For video telephony and conferencing, the delay requirement is about the same as for voice telephony and conferences. These are commonly indicated by MOS and R-values, ref. [Ulse06]. Regarding video a medium level of tolerance for errors in pixel and the corresponding bit representations can be assumed, since changes in no more than 1/1000 of the pixels on a screen are hardly perceptible. When coding and compression are applied, there is less failure tolerance in transmission, since corrupted or lost data can affect a large number of pixels after decoding at the receiver.

Video streaming does not impose strict timing requirements like the conversational applications with the exception of immediate delivery of live events. It is a one-way transmission from a server to the user or a broadcasting service for many users.

Traffic Rate [Mbit/s]:

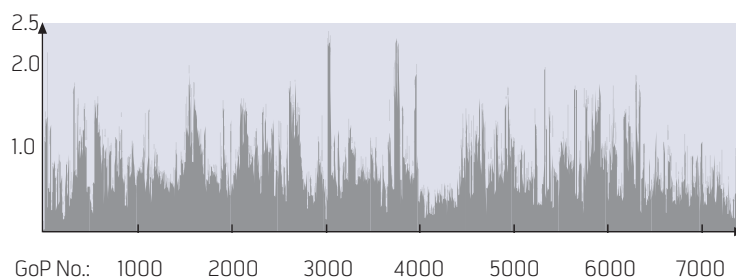


Figure 6 Variability of real time video traffic

Then the streaming traffic can be temporarily stored in buffers of usual size in routers and at the receiver. This in order to avoid data loss in short overload periods and for a play-out of video data independent of traffic fluctuation due to coding of frames at different rates within a group of pictures (GoP).

User experience is also influenced by the presentation device, such as HDTV screen, PC screen or mobile handset. This further complicates the challenge of estimated requested QoS levels.

There are multiple ways of implementing the transfer service, both in the network, but perhaps even more in the terminal equipment. For the latter, there are factors affecting the user perceived quality, like:

- Information coding applied;
- Packetisation efficiency, including how many information pieces are put into the same packet;
- Silence/no-change/no-activity suppression;
- Error-concealment methods;
- Codec-tandem performance.

One may say that the bit rate per session is a measure for the efficiency. Basically this efficiency can be increased by:

- Using low bit rate coding schemes;
- Increasing the packet lengths (less overhead compared to the payload);
- Multiplexing information from several sessions into the same set of packets;
- Compressing headers, e.g. for the combination of IP/UDP/RTP;
- Suppressing silence/no-change/no-activity periods.

The overall QoS requirements for data applications are related to impatience time as well as requirements on loss ratios. Commonly transport protocols and applications are able to detect loss of transfers, so these add more to the overall completion time. Still there are built-in timers, which often trigger re-transmissions, session initiation or session release.

Besides the “in-conversation” QoS parameters, user requirements are also present for session establishment and release. To some extent these may resemble the ones associated with classical telephone call handling. Parameters are defined by ITU-T Recommendation E.721 [E.721] as:

- Pre-selection delay
- Post-selection delay
- Answer signal delay
- Call release delay
- Probability of end-to-end blocking.

In the same recommendation, target values are also given, depending on the number of domains traversed, such as local, toll and international. For delay parameters the target values are given for mean values and the 95 % quantile. Most of these can be generalised for other application types as well.

Another important factor is the availability of transfer service. This does not only depend on links being operational, but also on all servers being on-line, such as the DNS server.

At an abstract level, similar parameters may also be described for data applications. One example is the time from clicking on a link until the first object of that page appears.

4 QoS mechanisms – system tools

4.1 Overall

In order to support transport of real-time traffic flows over an IP network, one must be able to handle:

- Timing and synchronisation of and between individual samples of traffic flows for the same applications;

- Effects of packets being lost;
- Effects of packets being delayed;
- Packets arriving in a different order at the receiver from how they were sent;
- Multiple traffic flows and different types of traffic flows;
- Monitoring and flow control.

As illustrated in Figure 7 a set of different mechanisms come to play when ensuring the service levels. Several of these are briefly described in this chapter. Note that some of these may be optional depending on the configuration and the overall solution for traffic handling.

Typically different mechanisms are used for the different network portions, such as access and core networks. On the border between network domains, border (or edge) capabilities are implemented. The services provided by a network domain refer to the phenomena observed at the border, also referred to as service demarcation point. Here, additional mechanisms are commonly implemented such as to monitor that the SLA conditions are met and alternatively enforce those conditions.

4.2 Traffic classes

Different regimes have been presented for defining traffic classes. One of these has been elaborated for UMTS (ref. [22.105]), as summarized in Table 1. These can be compared with the service types described above for voice, video and data. Four classes have been defined for UMTS:

- Conversational – connecting peers or groups of human end-users who expect an unnoticeable low delay and thus have strict real time demands;

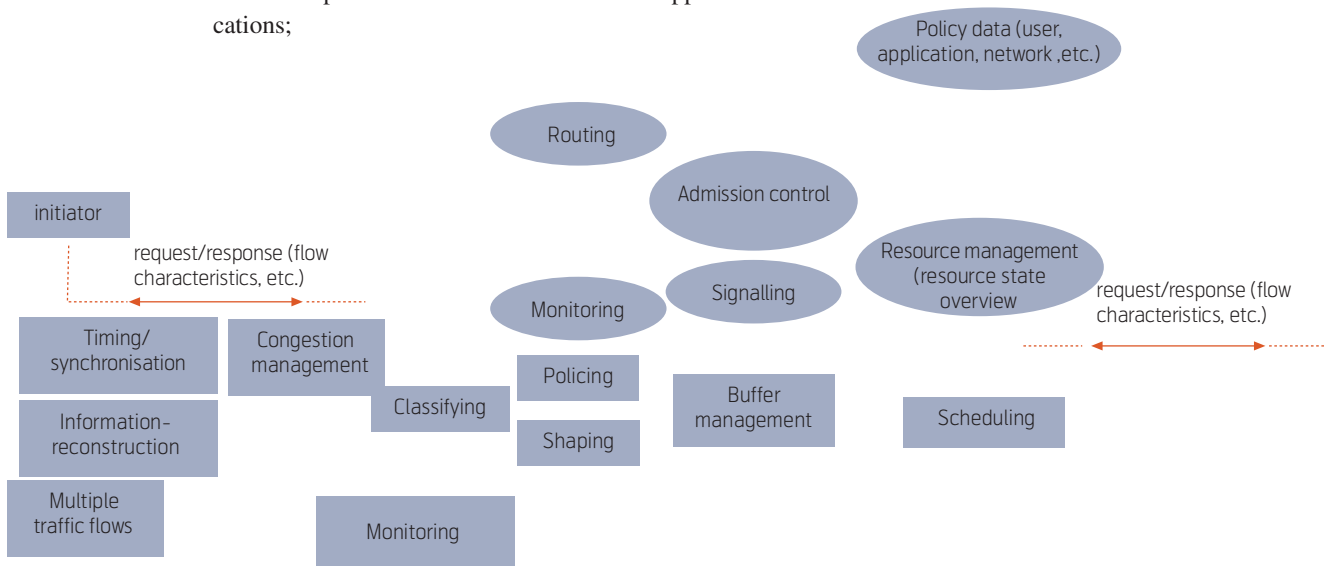


Figure 7 Illustration of features related to service level guarantees (by the network)

Traffic class	Conversational	Streaming	Interactive	Background
Quality of service demands	Stringent and low delay; Relaxed error and noise tolerance	Preserve time relation (variation) between information entities	Preserve payload content; Relaxed timing constraints	Preserve payload content; Delay tolerance for hours or days
Traffic pattern & environment	Bidirectional or multicast real time data delivery	One way or broadcast continuous streaming	Request response pattern for information retrieval	Automated data download without human interaction
Applications	Telephony (GSM, UMTS, VoIP), Video conference	Video on demand, Internet radio & television	Web browsing, human communication to servers; polling; distributed processes	Peer-to-peer downloads; file transfer; email; SMS

Table 1 Traffic and service classes defined for UMTS [22.105]

- ii) Streaming – intended for one-way transport of voice, audio or video data which is listened to or looked at by a human at the destination. The delay constraint is not strictly real time but may differ in the range of some seconds to minutes, but the time relation of the video or audio has to be preserved from the sender to the receiver side;
- iii) Interactive – assumes machine or human end users to request data from remote equipment as a server. The class does not introduce real time constraints but expects data to be delivered error-free or at a negligible error rate;
- iv) Background – addresses transfers between machines with large delay tolerance allowing for bandwidth reduction or interruption in favour of other traffic at high load. Again, the data transfer is expected in assured mode avoiding errors in the data.

In principle, the QoS requirements from the users' view are the same regardless of whether communication systems are wired or wireless. However, several factors may lead to different expectations or perceptions of the service level. Examples are noise in the user environment, awareness of technical restrictions and capabilities of the end-device.

Indications of performance targets are given in several sources, such as [G.1010] and [22.105]. Note that different assumptions may well be laid down, resulting in different values. Typical services are audio, video and data. The different target values inspire for realising a set of service classes in an IP-based network. Differentiated Services (DiffServ) is promoted as a service architecture supporting a scaleable way to achieve relative service and QoS levels in an IP network. DiffServ operates on aggregated flows by dividing the traffic flows into a set of classes.

DiffServ uses a particular implementation of the IP version 4 Type of Service (ToS) header field. This field is now called the DiffServ field, consisting of 8 bits, out of which 6 bits are available for current use and two are reserved for future use. The 6 bits define the DiffServ Code Point (DSCP), which identifies a Per-Hop Behaviour (PHB). The PHB indicates the way packets shall be handled in the routers and can be set and reset in any DiffServ capable node, also referred to as marking the IP packet.

ITU-T Recommendation Y.1541 outlines a potential mapping between IP QoS classes and the IP DiffServ classes. An abstract of this is given in Table 2. As commercial IP-based networks have evolved to support guarantees lower than the values shown, it would be expected that more stringent classes are offered.

Traffic types from different applications would then be related with the service classes. However, in several commercial cases, the user category is also taken into account when this mapping is done. In this manner different voice classes, for example, could be defined, e.g. depending on the pricing of the different voice service classes. Therefore, both application requirement and user group are looked at (Figure 8).

One example is illustrated in Figure 9, where two user groups are present: i) residential users, and, ii) visiting users. The distinction between the residential and visiting users is essential for the Open Access Network, OAN (ref. [Øste05]) concept where the main idea is to offer excess capacity to visiting users and where it is assumed that the visitor only to a small extent shall influence the network performance of the residential user.

Due to the access technology currently available, the possible bottlenecks in the typical OAN configuration are foreseen at mainly two places in the network:

IP DiffServ class	Y.1541 IP QoS class	IP packet transfer delay	IP packet delay variation	IP packet loss ratio	IP packet error ratio
EF, Expedited Forwarding	Class 0	100 ms	50 ms	10^{-3}	10^{-4}
	Class 1	400 ms	50 ms	10^{-3}	10^{-4}
AF, Assured Forwarding	Class 2	100 ms	Unspecified	10^{-3}	10^{-4}
	Class 3	400 ms	Unspecified	10^{-3}	10^{-4}
	Class 4	1 s	Unspecified	10^{-3}	10^{-4}
Default	Class 5	Unspecified	Unspecified	Unspecified	Unspecified

Table 2 Mapping IP DiffServ classes and ITU-T Recommendation Y.1541 IP QoS classes (note that several of the values are under further study)

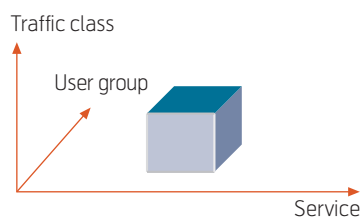


Figure 8 Identifying traffic class by considering both service/application and user group

- Radio link between UE (User Equipment) and RGW (Residential Gateway);
- Access line between RGW and the public network.

A main challenge in the OAN concept is that both the residential and visiting users will compete for rather limited WLAN resources as well as the capacity of a rather slow access line (xDSL).

4.3 Differentiation

Having defined a set of *traffic classes* allows for *utilising differentiation mechanisms*. The overall objective of differentiation is to *achieve higher resource utilisation* as traffic classes with stricter requirements get preferential treatment. Still requirements of all

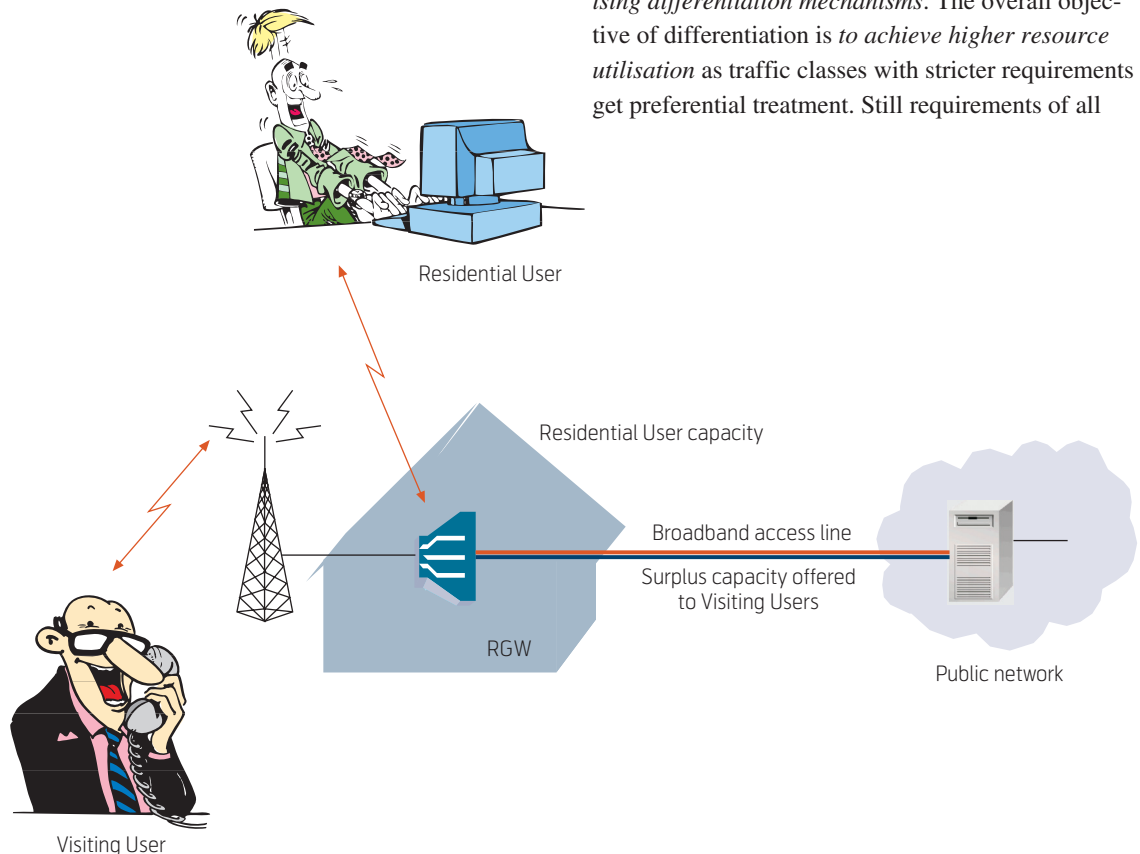


Figure 9 The Open Access Network Concept of making surplus capacity of residential access subscriber lines available to (casually passing) visiting users

traffic classes should be met. Differentiation can be done according to several parameters, three being packet delay, packet loss and service dependability. Only the first is discussed here.

The situation arising on the access line is illustrated in Figure 10. The different traffic classes may be mapped into service classes as they are recognised by the nodes at the two ends of the access line. For the residential these nodes would commonly be the residential gateway and the service edge router (edge router). Different traffic classes may also be defined in the core network and in the customer network.

As depicted in Figure 11, three basic operations must be done for efficient differentiation on delay to take place: i) classifying packets, ii) buffering packets into proper queues, and, iii) scheduling individual packets for transmission.

In order to achieve differentiation in practice, the proper mechanisms have to be activated and the load has to be controlled. As the load varies during the day, this may place additional requirements on the traffic handling mechanisms.

Referring to the OAN concept, see Figure 9, there are some important questions to be raised for multiplexing on the xDSL line. If we take the current capacity for typical xDSL the upstream capacity is limited to less than one Mbit/s. This rather limited capacity has

to be shared between both the residential user and the visiting users. The performance of the xDSL part will therefore have a large impact on the QoS as seen by an OAN user. One possible approach to differentiate among users could be to apply the DiffServ classification model and simply make the QoS differentiation on the IP level, i.e. the scheduling is done completely on the IP level. This approach may have some negative implications for the QoS for real time traffic due to the fact that the delay and jitter may increase because of highly variable sizes of the IP packets for real time traffic and data (elastic) traffic.

It should be emphasised that the QoS perspective for a service is an end-to-end basis, so the QoS seen by a user (end-to-end) will to a large extent be determined by the performance of the link (or network element) that appears to be the bottleneck. In the OAN concept with a well-dimensioned core, the presumable bottlenecks will either be the WLAN part or the access line (xDSL).

The effect of introducing several priority classes for QoS differentiation is not obviously effective especially if two classes have more or less the same requirements. With two classes, where the higher one is for real time traffic and the lower for elastic traffic, it is possible to give guarantees for delay quantiles by keeping the corresponding load strictly less than an upper limit. To divide the capacity between residential and visitor the loads from each of these classes

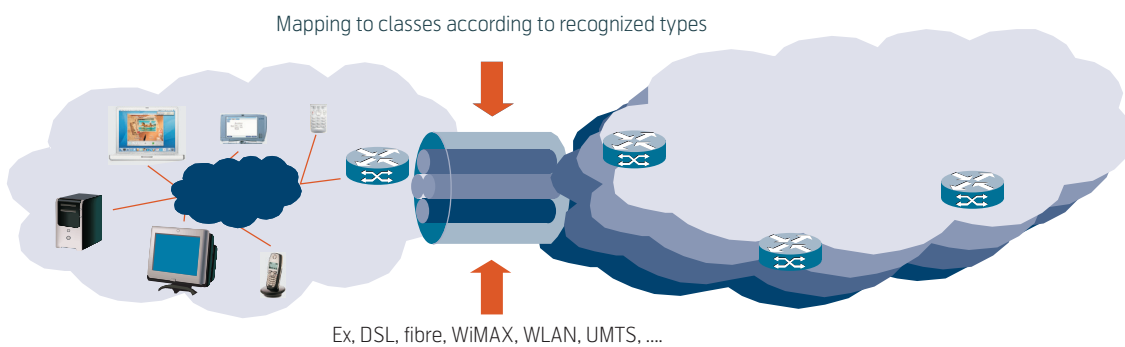


Figure 10 Organising a number of classes on the access line

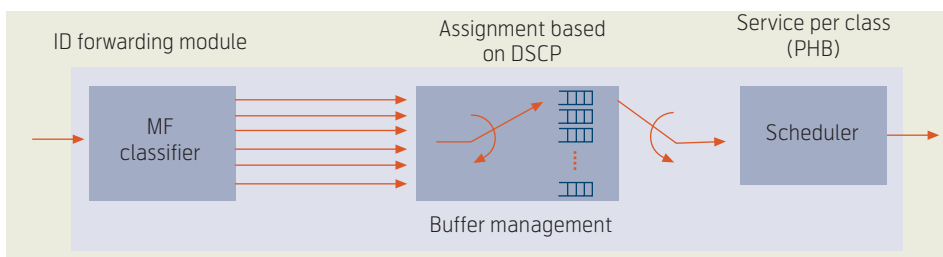


Figure 11 Functional components for supporting differentiation

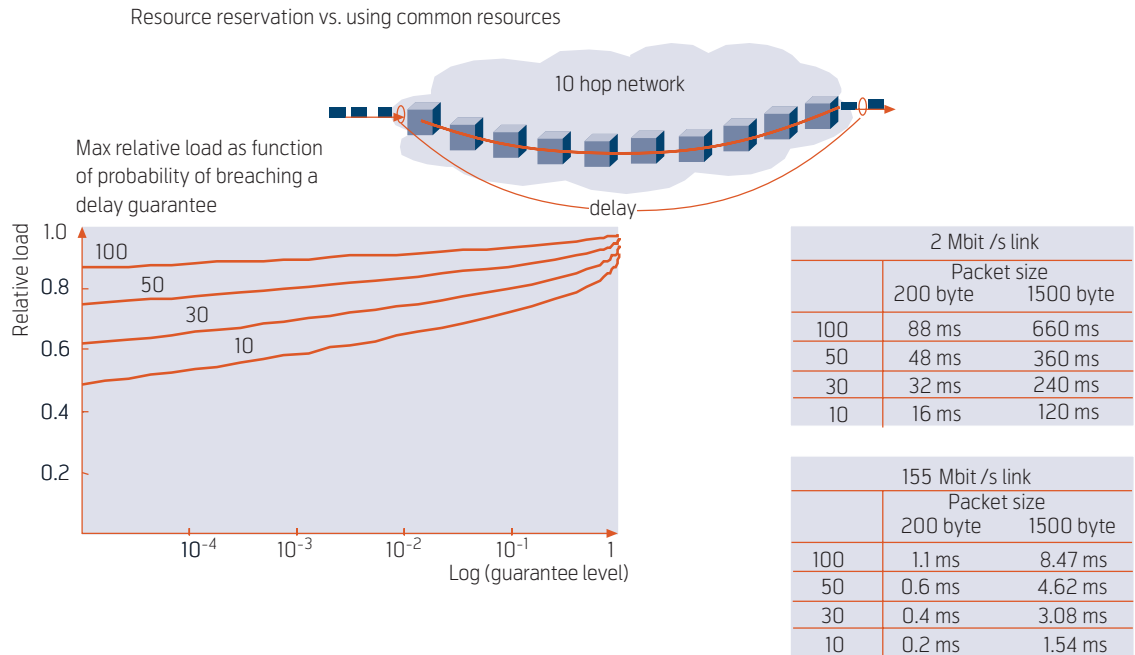


Figure 12 Illustration of service level guarantee that can be used for a 10 hop path

have to be limited as well as the sum. It turns out that a three level scheduling scheme gives delay quantiles for the medium priority traffic that is more sensitive to load variations in the high priority traffic. However, by controlling the high and medium priority loads in the range 0.3 – 0.4, the difference between a two-level and a three-level scheduling configuration is limited.

As traffic of different types is multiplexed on a link, this may cause delay and jitter issues. One cause for jitter is the variation in the packet lengths for the different traffic types. While typical real time traffic like voice will emit packets of a small fixed size, the typical data application may generate packets that are quite long. Due to this variation in packet size between different applications the queueing delay for typical real time applications may increase over the limit where degradation is inevitable. This negative multiplexing effect will add for each router along the path from the sender to the receiver. However, for high capacity links this queueing delay will be more or less negligible, leaving the main delay contribution to low capacity links in the access network.

The main observations are that IP multiplexing for high capacity links such as STM-4 (622 Mbit/s) or higher should not cause any particular problems for real time traffic such as telephony. The jitter is well limited below 2.0 ms for up to 13 hops. As a result one may conclude that the buffers needed to restore a constant bit stream in the gateways could be limited to only a few ms. The same result will also be valid

for STM-1 (155 Mbit/s) links but with a minor increase in the jitter (as seen from Figure 12). However, it should be mentioned that internal capacity limitations inside the gateways could limit the throughput and can therefore cause some additional queueing delay, for instance due to limited processor capacity. If a 2 Mbit/s link, however, is used all along the path, jitter may be significant. This is seen from the graph in Figure 12 showing maximum relative load (between 0 and 1 – vertical axis) allowed on the link in order to avoid a guarantee level (up to 1 – horizontal axis) being exceeded for different delay values.

Another illustration is found in ITU-T Recommendation Y.1541, see Figure 13. This shows the allowed link load of the delay-sensitive traffic flow as a function of number of hops for requirements on different delay quantiles.

The terms in the Service Level Specification (SLS) are checked at the border of the domain, e.g. in an edge router. In case the appropriate DSCP value has not been inserted in the packet, this has to be done, based on various combinations of information in the packet. This information may include the IP packet header as well as the header of the transport protocol. Additional information may also be used, like the interface on which the packet arrived on. This means that a Multi-Field (MF) classifier and marker would be activated in the first DiffServ-capable router in the domain. Then, the packet has got its DSCP. Characteristics of a flow (aggregate) will then be monitored to see whether the packet is forwarded directly (con-

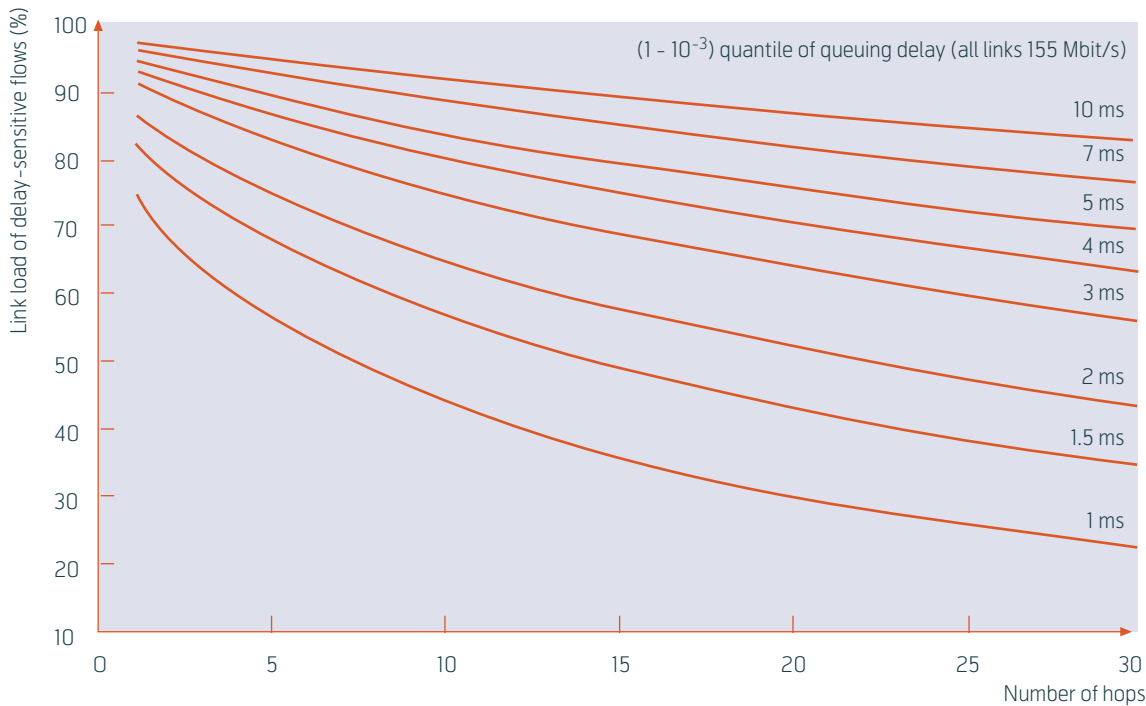


Figure 13 The $(1 - 10^{-3})$ quantile of the overall delay for different levels of variation-sensitive traffic (vertical axis) and for different numbers of router hops (horizontal axis), from [Y.1541]

ditions in the SLS are obeyed), being dropped, re-marked or shaped. A logical relation between the different functions is illustrated in Figure 14.

Assuming that a traffic mixture is present in the network, i.e. traffic flows having different requirements on delay/jitter, it can be shown that having mechanisms for separate treatment of the traffic classes allows for better utilisation of the link capacities, compared to when only a single class is supported. This is at the expense of fairly increased processing complexity. The software/buffering/processing cost is balanced against the link cost. The termination boards should also be counted into the link cost. In addition to the functional blocks, the traffic handling must be able to operate on individual DiffServ groups, like buffering/queueing discipline, and so forth.

Having a traffic mixture, e.g. voice and web surfing, the traffic class with the tightest performance requirements may decide the link/bandwidth needed. For example, the delay and jitter requirements for the voice traffic may limit the utilisation of a link. This is illustrated in Figure 15. Two traffic classes are assumed; class 1 having a much stricter requirement on delay than class 2. As the load on the link increases, the delay restriction on class 1 is faced, say at load ρ_s . In case the two classes are separated, a higher link load is accepted before the delay restriction is faced.

Hence, when DiffServ is introduced, the traffic classes can be separated. This may allow for an increased link utilisation compared to best effort, while still meeting the delay/jitter requirements for the class with the tightest performance requirements.

4.4 Admission control

Admission control is a *preventive traffic control that aims to admit an arriving new traffic source if and only if its quality of service as well as that of the already accepted sources is guaranteed. The admission control procedure should also ensure a high utilisation of network resources through efficient statistical multiplexing.*

Running an application, a number of flows might result. An example is a multimedia application covering voice, video, file transfers, interaction control, etc. All these flows should be served in order for the

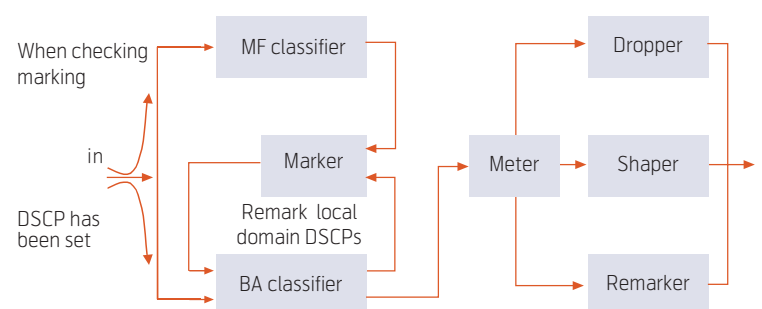


Figure 14 Logical view of traffic conditioning in a DiffServ node

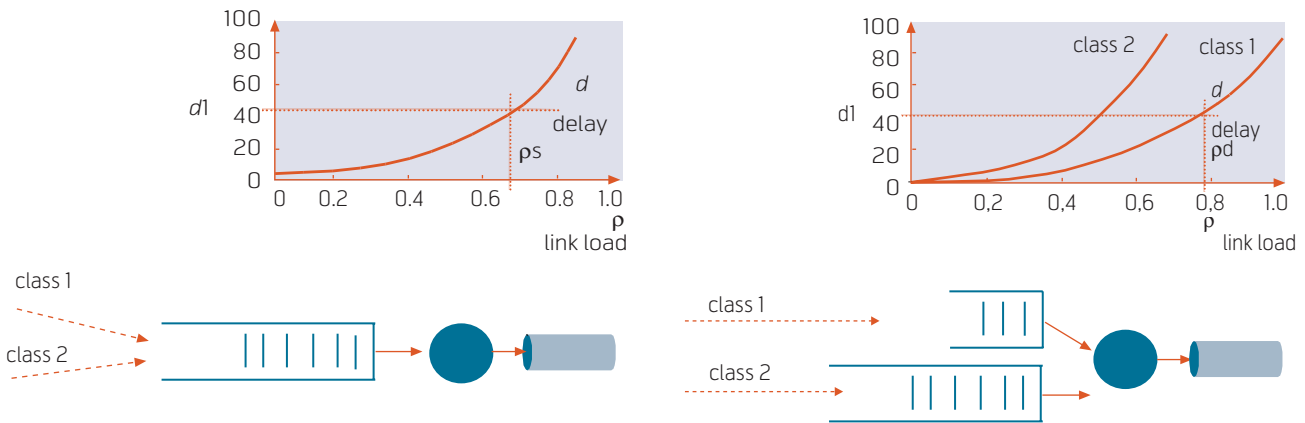


Figure 15 Illustration of increased link utilization by differentiation between traffic flows

application to be run in a satisfactory manner. Hence, the term session is introduced. A *session* is a continuous period of activity during which a user generates a set of flows (elastic or streaming type).

It should be noted that the session term is seen with varying interpretations, like a SIP session, an FTP session, an HTTP session, and so forth. This relates to media supported (voice, video, data). Usually, the admission control acts on the flow level, not taking into account effects on the session level. An argument may be that an application, in case a flow is not accepted, may retry the transfer and then leave that operation to the end-system/application. Hence, the network would not need to be enhanced with capabilities enabling grouping of flows into sessions. For some services and users, however, the service portfolio (and the SLA conditions) may refer to phenomena on the session level. This is not covered here as any correlation between those levels might be estimated by monitoring/measuring, e.g. for verifying the SLA conditions.

The overall objectives of having admission control are to:

- Ensure that the existing traffic flows still receive adequate service levels when additional traffic flows are introduced;
- Provide appropriate feedback/advise to a user/application when initiating a session that this session (or traffic flow) may well receive too low service performance;
- Enable differentiation between traffic flows, including applications and users in accordance with policy and subscription/user profile;

- Balance ensured service provision (with effective guarantees on performance levels) and efficient utilisation of network resources.

These objectives will not be equally weighed independent of the scope of discussion. For instance, from a user perspective, less interest could be placed on the network utilisation issue.

Integrating elastic and real-time traffic on the same resource units may allow for increased efficiency. By giving priority to the real-time flows, they could experience a resource that is almost loaded as if they were the only active flows. Then, elastic flows could be served whenever the resource is not used by the real-time flows. However, this may introduce long delays for the elastic flows during some periods. An approach is to restrict the load from the real-time flows, ensuring that some capacity is available for the elastic flows.

As mentioned earlier, executing the admission control algorithm would basically provide an answer whether to accept or reject a request for serving a traffic flow. In principle the answer could also be to accept but on certain conditions, like some characteristics of the existing flows having to be changed/renegotiated.

Then, in order for the algorithm to arrive at that decision, a number of inputs have to be available. Hence, the algorithms may differ in terms of which inputs that are needed/taken into account. Furthermore, the algorithms may also differ in terms of which answers that are possible (only accept or reject, or more subtle outputs). An illustration of inputs and outputs is given in Figure 16.

For the admission algorithm various scopes and principles could be taken, like:

Implementing admission control

RSVP-based

As a general signalling protocol, RSVP may carry most of the data needed for admission control, including characteristics of the traffic flow as well as information about the users/port numbers.

Initiating the RSVP messages by the end-systems, the traffic handling mechanisms may be co-ordinated dynamically along the relevant data path. In some places this is referred to as dynamic topology-aware admission control.

RSVP is used by an end-system to request specific service levels from the network for particular traffic flows. Routers also apply RSVP to forward requests to all nodes along the path(s) of the flows and to establish and maintain state to provide the requested service. Hence, RSVP requests will generally result in resources being reserved in each node along the path. RSVP allows users to obtain preferential access to network resources under the control of an admission control mechanism. Such admission control is often based on user or application identity, however, it is also valuable for providing the ability for per-session admission control. In order to allow for per-session admission control, it is necessary to provide a mechanism for ensuring that an RSVP request from an end-system has been properly authorized before allowing the reservation of resources. In order to meet this requirement, there must be information in the RSVP message which may be used to verify the validity of the RSVP request. An example is to have an authorization element assigned to the user, which can be inserted in the RSVP messages.

Policy and Bandwidth Broker-based

Policy and Bandwidth Broker (BB) are described in later sections. Although RSVP supports the ability to convey requests allowing for resource reservations, an essential feature may be missing. This feature is the ability of network managers and service providers to monitor, control, and enforce the use of network resources and services based on policies derived from criteria such as the identity of users and applications, traffic/bandwidth requirements, security considerations, and time of day/week. A framework for policy-based control over admission control is described in [RFC2753].

Implicit admission control / Local probing

The local probing approach relies on generating test packets (probes) to check whether or not a new traffic flow can be set up. The probes may be generated by the end-systems. In case several service classes are offered, it is to be decided if the probes should be sent in the same class as the following traffic flow or in another class (e.g. the lower service class). Hence, the local probing may be suitable for DiffServ. An advantage of this method is that no changes are needed in the routers not generating probes. This has also been referred to as distributed admission control, see e.g. [Kell00].

Probing results may also be based on marking (e.g. using ECN) or monitoring errored packet arrivals (e.g. using RTP/RTCP) of ongoing traffic flows. Hence, information on the marking tells the admission control algorithm whether or not a new traffic flow with certain characteristics can be served.

A common feature with implicit admission control is that no per-flow state information is needed, which also may be run in the end-systems. However, remembering the connectionless nature of IP, and if the routers are not taking part in the control, it may be uncertain if all packets in the traffic flow actually traverse the same path. This means that some mid-flow packets may well experience other conditions than the information estimated from probes if they are transferred on another path.

The different schemes for admission control may be combined. For example, an implicit admission control may be used in the access network (between the terminal/host and the edge router) while other schemes are used in other parts of the network.

- Which time scale is considered: Is only the current situation taken into account or is a more future-oriented approach followed? Is the decision to be based also on historic/trend information? An example is that a traffic flow accompanied by low revenue could be rejected even if sufficient capacity is available if there is a high probability that a flow accompanied by higher revenue had to be rejected later on.
- What level of “gambling” is used for guaranteeing the service level? Rather loose thresholds have to

be used if strict guarantees are given, while tighter thresholds (and even overbooking) could be used when a more “gambling”-like attitude is taken on.

4.5 Resource reservation, dependability/ resilience differentiation

An idea behind the *principle of capacity reservation and controlling traffic load is to avoid SLA conditions / QoS expectations being violated*. Breaching an SLA condition may ask for exercising discount/refund schemes, reputation impairments, and so forth. That is, this can be considered as an expense side to

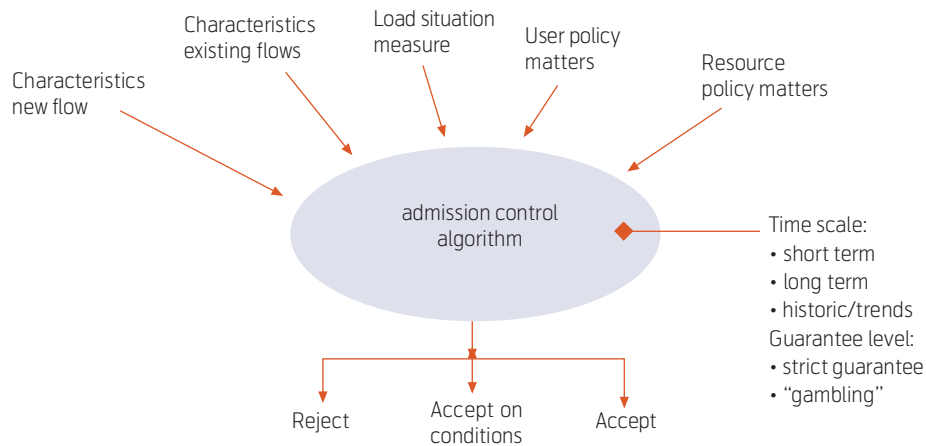


Figure 16 Input and output candidates for the admission control algorithm

be avoided, which can be balanced against the increased complexity needed. In principle, all these aspects could be quantified in order to examine to what extent introducing these mechanisms would be worthwhile.

When reserving capacity for a flow (or flow aggregate) a setup/signalling protocol can be used. One option is to apply management-related procedures for this purpose, implying that the management system could interact with the routers instead of signalling being exchanged directly between routers. In addition, a combination of management and signalling procedures may also work, for example when combining the action with policy matters, bandwidth brokers, and so forth.

The so-called scalability issue implies that reservation for individual flows is not realistic in the core network where lots of flows are present. On the other hand, for the access line, resources could be reserved, e.g. during an active session.

In a network different resource types may exist, including processors and links. These could well be set up with different dependability schemes. On the

node/processor level, load-sharing, hot-standby, common spare and so forth, are applied. That is, configurations such as $N + 1$, $1 + 1$, $1 : 1$, etc. are used. A key feature is to define triggers for when a switch-over should take place, that is, which events can take place before a resource unit is declared as unusable.

Regarding links there are commonly a set of layers, say running IP over MPLS over an optical layer, ref. [Jens03]. The dependability has to be tuned across these layers for an efficient operation. Some examples of these options are outlined in Table 3.

The characteristics of the different recovery models are:

- Protection switching: The alternative path is pre-established and pre-reserved (pre-provisioned). Hence, the shortest traffic disruption is achieved. Two main groups are $1 + 1$ and $1 : 1$. In the former packets are forwarded simultaneously on working and protection path. When the working path fails, the downstream node simply selects packets from the alternative path. For $1 : 1$, packets are forwarded on a predefined path in case of failure on the working path. When no failure is present, the alternative path may carry other traffic flows.

Recovery model	Protection switching		Restoration (MPLS rerouting)		(IP) rerouting
Resource allocation	Pre-reserved			Reserved on-demand	
Resource use	Dedicated resources		Shared resources		Extra traffic allowed
Path setup	Pre-established		Pre-qualified		Established on-demand
Recovery scope	Local repair	Global repair	Alternate egress pair	Multi-layer repair	Conc. prot. domain
Recovery trigger	Automatic inputs (internal signals)			External commands (OAM signalling)	

Table 3 Examples of recovery options

However, these flows must be pre-emptible as they should be dropped if the alternative path is needed for the protected traffic.

- Restoration (MPLS rerouting): The recovery path is established on-demand after detecting a failure. As it takes a while to calculate new routes, signal them and configure the relevant mechanisms, this may take considerably longer than protection switching.
- IP rerouting: Ordinary routing protocol and exchange principles are utilised for identifying alternative paths.

These dependability mechanisms are introduced to assist in complying with the service levels. In addition to the basic up and down states of a resource, differentiation during failure situation is an area of much interest. Differentiation criteria could be service types like video, voice and data, or customer types. One way is to map these service types into logical trunks, e.g. MPLS paths. These could then be given different resilience schemes.

4.6 Monitoring

Managing any network or service provision, defining and following an adequate set of performance indicators is a necessity. Such indicators are typically used in order to assess the “health condition” of the operation and service delivery. In general both technical and financial indicators will be used, in addition to others, e.g. reputation. However, technical-related ones are the main topic here. Again, these could be divided into separate parts, for example referring to different portions of a system and different phases of the service provision.

Monitoring traffic flows, resource utilisation and service levels has been an activity for quite a few decades. Still there seems to be some striving to find the proper balance between achieving an adequate picture of conditions in the system and not spending too much resources on monitoring. One centralised approach is to monitor servers and common network resources. This may save some monitoring equipment, although too many averaging operations might hide problematic portions. A fully distributed approach is to have monitoring agents installed in user devices, although a management challenge would then be seen together with the “trust level” between the user and the provider.

Considering the multi-service, multi-technology, multi-provider situation, the monitoring challenge will grow further. A specific objective is to apply the monitoring results to trigger certain actions, either

by the operator/provider or by the user. Multiple purposes could be defined, both on enhancing the capacity (or re-configuring the available capacity) or restricting the traffic load (admission control, policing, charging, etc.).

- In order to reach a situation where users and providers of IP services have a common understanding of performance of the network, a set of harmonised IP performance metrics has been devised.

A series of RFCs has been issued for specific performance metrics:

- RFC 2330 Framework for IP Performance Metrics, IPPM
- RFC 2678 IPPM Metrics for Measuring Connectivity
- RFC 2679 A One-Way Delay Metric for IPPM
- RFC 2680 A One-Way Packet Loss Metric for IPPM
- RFC 2681 A Round-Trip Delay Metric for IPPM

Performance parameters related to forwarding of IP packets have also been described in ITU-T, see e.g. [Y.1540] and [Y.1541]. These include IP packet delay variation, IP packet error ratio, IP packet loss ratio, IP packet transfer reference event, IP packet throughput, IP packet transfer delay, and spurious packet ratio.

The Real-time Traffic Flow Measurement (RTFM) Working Group has described a measurement architecture to provide a method for gathering traffic flow information, see [RFC2722]. The model proposed is based on the concepts of meters and traffic flows given as:

- Meters observe packets as they pass by a single point on their way through the network and classify them into certain groups. For each such group a meter will accumulate certain attributes (such as number of packets and bytes). These metered traffic groups may correspond to a user, a host system, a network, a particular transport address (e.g. a port), etc. Meters are placed at measurement points and selectively record network activity as directed by its configuration settings. Meters can also aggregate, transform and further process the recorded activity before the data is stored.
- Traffic flow is said to be a logical entity equivalent to a call or connection. A flow is a portion of traffic that belongs to one of the metered traffic groups mentioned above. Attribute values (source/destination addresses, number of packets, etc.) associated with a flow are aggregate quantities reflecting

events that take place. Flows are stored in the meter's flow table.

A traffic meter has a set of rules which specify the flows of interest. One way to identify a flow is by stating values of its address attributes.

As well as flows and meters, the traffic model measurement includes managers (to configure and control meters), meter readers (to transport recorded data from meter to analysis applications), and analysis applications (to process the data from meter readings so as to produce whatever reports are required).

4.7 Policy management and control

Emerging trends include policy-based implementations as observed in documents by 3GPP, ETSI-TISPAN and others. Then, policy control and execution points are defined. This could support the dynamics by adapting policy rules for traffic handling. The policy rules state which network/traffic situations are considered when appropriate actions are defined. This also supports user/application-dependent rulings. The two-stage set-up process would become the core in this concept. A result is that admission control and SLA conditions can be followed with greater granularity.

Policy can be considered as a set of principles for usage of resources, given by business considerations. That is, the business decisions are translated into statements relevant for the usage of resources in the network.

The semantics of a *policy rule* is a conditional imperative statement in the form

```
if <condition> then <action>
```

Thus, applying a rule means to evaluate its condition (matching the rule), and, depending on the outcome of that, either execute the action or not. Policy rules may be nested.

Policy-based network management would provide a centralised platform for network managers for defining and distributing network policies to enforcement points throughout a network. In a typical policy-based framework, see Figure 17, the network manager edits policies through a policy entry console. Those policies are then stored in a policy repository. When requested, a policy server (Policy Decision Point) retrieves policies from the repository and makes policy decisions that are communicated, e.g. applying Common Open Policy Service (COPS), to the relevant network points. Those network points, like routers, switches and firewalls, enforce the policy decisions in the network. COPS is a query and response TCP-based protocol that could be used for exchanging information between Policy Decision Point (PDP) and Policy Enforcement Point (PEP).

An example of a policy is

- i) if "traffic flow within profile X" then "mark packet with DSCP = AF1";
- ii) if "traffic flow out of profile X and within profile Y" then "mark packet with DSCP = AF2";
- iii) if "traffic flow out of profile Y" then "drop packet";

where profile X can mean rate is less or equal to 64 kbit/s, and profile Y can mean rate is less or equal to 128 kbit/s.

The IETF Policy Working Group standardises the basic framework of policy-based management system for IP networks. It focuses on representing, managing and sharing policies in a vendor independent, interoperable and scalable manner.

The Policy WG co-ordinates the development of the QoS schema with the Policy Information Base (PIB) and the Management Information Base (MIB) being developed in the DiffServ WG as well as with extensions to the COPS being developed by the Resource Allocation Protocol (RAP) WG.

Policy rules must be represented as data structures so they can be stored and retrieved. To address this issue, the IETF Policy Working Group has defined the Policy Framework Core Information Model, which defines a high-level set of object-oriented classes that can be used for general policy representation.

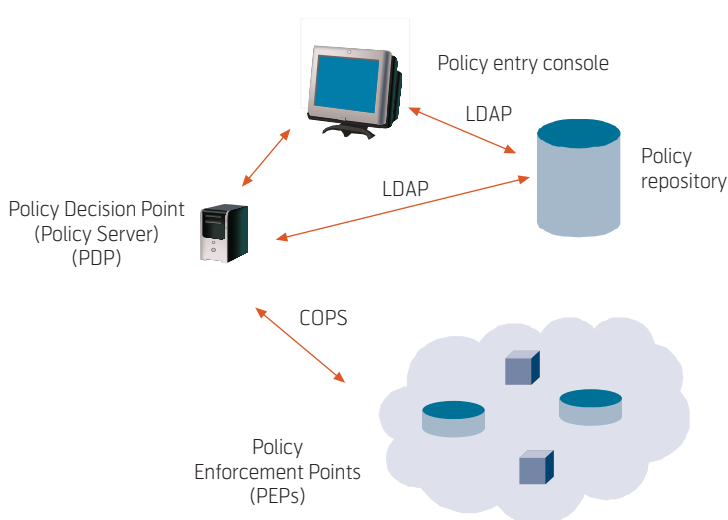


Figure 17 Example of Policy Framework components and protocols

5 Interconnecting domains

Even within the administration of one operator a number of domains may be defined. One reason for this can be that different flow granularities are requested. For example, on the access link of a household the capacity is commonly restricted and therefore performance and even admission guidelines on individual flows can be implemented. This would not be feasible in a core network where traffic related to millions of users is flowing.

An access link may thereby be supported by resource reservation, policing, packet priorities and other features as described in the previous chapter. In a core network DiffServ, in combination with MPLS, would likely be implemented. Combining a tighter regime and DiffServ can bring some benefits compared to “simple” DiffServ. One example is that admission control can be applied at the border of the DiffServ domain. Explicit signalling per flow allows for admission control such that the flows in each traffic class receive the service level expected. Voice conversations are examples where admission control could be fruitful to ensure that the ongoing conversations get the service level and additional conversations are rejected in case there are not sufficient network resources.

Schematically, a multi-domain path can be illustrated as in Figure 18. Note that this shows a point-to-point configuration, also referred to as a “pipe model” where the two end-points are given. Such an illustration could be used for defining a hypothetical reference connection. This can be assumed for allocating service level degradation to the different network segments. On the other hand, the allowed degradation may also be negotiated pair-wise between the providers in the chain on commercial terms.

The different QoS parameters can be accumulated along a path in different manners. For example, a mean delay parameter is additive; contribution from each segment can be added. For calculating loss ratio, error ratio and availability, a multiplication is commonly applied. Delay variation is often estimated by convoluting, see e.g. ITU-T Recommendation Y.1541 amd. 2.

When explicit signalling per flow is used, policy-based control, e.g. per user and per application can be introduced in a more dynamic way. Moreover, if the router in the network is doing the packet marking signalling can be used to convey the information to the router on which DiffServ class to apply for each flow. This would particularly be useful in case IPSec is applied if the IP addresses and port numbers are not statically assigned to DiffServ classes.

5.1 User's rights – AAA functionality

When providing a service commercially and different service levels are used, they have to be supported by corresponding AAA (Authentication, Authorisation, Accounting) functionality. Accounting can be seen as included in the service provisioning process (called integrated accounting) or it can be offered as a separate service (called discrete accounting). In the former the accounting is coupled to a specific service, collecting relevant information by using service specific entities.

For getting access to a service, the user sends a service request to the AAA server. This checks the authorisation of the user and, assuming access is granted, forwards the necessary information to the relevant server. This server finds the information relevant for configuration of the network resources (service equipment) and distributes this information

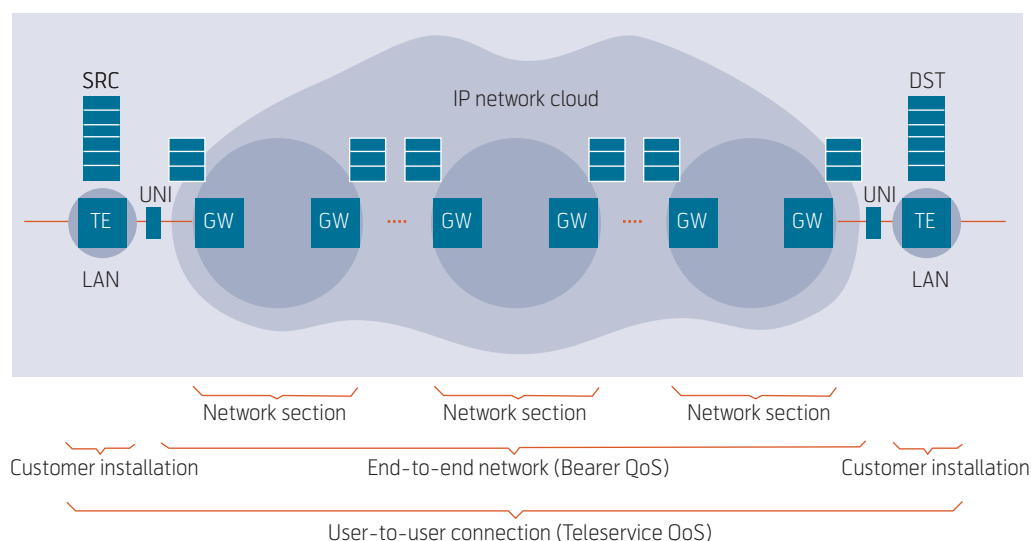


Figure 18 Reference path illustration from ITU-T Recommendation Y.1541

to the network nodes. In case of DiffServ, the accounting system, QoS control and Bandwidth Broker are noted.

In ETSI-TISPAN, these actions are initiated by the Network Attachment Subsystem, NASS, ref. [ETSI-282001]. NASS provides the following functionalities:

- Dynamic provisioning of IP addresses and other terminal configuration parameters;
- Authentication taking place at the IP layer, prior to or during the address allocation procedure;
- Authorisation of network access based on user profiles;
- Access network configuration based on user profiles;
- Location management taking place at the IP layer.

In order to carry out its tasks, the NASS interacts with the HSS, which is a core element for storing user data, ref. [Ross06].

5.2 Managing resources

Returning to the ETSI-TISPAN reference architecture, the resource and admission control functionality (RACS) contains mechanisms for

- Admission control, including checking authorization based on user profiles held in the access network attachment subsystem, on operator specific policy rules and on resource availability;
- Gate control, including network address and port translation, priority packet marking.

In the same architecture, a Border Gateway Function (BGF) provides the interface between two IP-transport domains. A BGF may reside at the boundary between an access network and the customer premises equipment, between an access network and a core network or between two core networks. It supports one or more of the following functionalities:

- Opening and closing gates (i.e. packets filtering depending on “IP address / port”);
- Allocation and translation of IP addresses and port numbers (NAPT);
- Interworking between IPv4 and IPv6 networks (NAPT-PT);
- Topology hiding;
- Hosted NAT traversal;
- Packet marking for outgoing traffic;

- Resource allocation and bandwidth reservation for upstream and downstream traffic;
- Policing of incoming traffic;
- Antispoofing of IP addresses;
- Usage metering.

As noted above the RACS applies a policy concept, see Section 4.7. The components are similar to the ones depicted in Figure 17. The Bandwidth Broker (BB) concept is similar to a Policy Decision Point (PDP) in the sense that it makes decisions regarding bandwidth provisioning. However, bandwidth brokers tend to operate at a higher level than PDPs. PDPs are typically connected to a (small) number of PEPs within an administrative domain. They tend to be topology-aware as a result of their role, e.g. in the RSVP admission control process. Bandwidth Brokers are aimed more at the interfaces between domains.

A BB refers to an abstraction that automates the admission control decisions for service requests in a network domain. This means that it is responsible for keeping track of the current allocation of reserved traffic, it is configured with policies that define which traffic flows belong to which traffic classes, and it interprets new requests in the light of these policies and the current bandwidth usage. In this sense, a BB can be considered as a special type of policy server that is responsible for those related policies for a network domain. A BB is not necessarily a policy manager but policy management and bandwidth brokering will need to work together in providing integrated policy services and admission control. Another important function of a BB is to configure network devices according to admitted QoS requests.

In the intra-domain case, the BB manages the resources based on the SLA/SLS that has been agreed upon between domains. One or more protocols are used to exchange information between a host and a BB, and a BB and a router.

In the inter-domain case, the BB is also responsible for managing inter-domain communication with BBs in neighbouring networks. This is to co-ordinate SLAs across boundaries. In order to co-ordinate bandwidth assignments across domains, a single inter-domain BB protocol must exist.

Figure 19 shows a sample network configuration. It consists of three domains AS1, AS2, and AS3 with a BB for each one (BB1, BB2 and BB3). The SLAs/SLSs are placed between AS1 and AS2, and between AS2 and AS3. A user can be either an end-system or an application that requests bandwidth.

The Bandwidth Broker makes decisions based on the network topology and the network traffic characteris-

tics. The network topology consists of a description of all available network resources: nodes, links, link metrics, physical link capacities, e.g. link capacity that can be allocated, resource class (gold links, links only to be used for premium customers), etc. The network traffic characteristics are expressed as a set of traffic trunks which mainly express a bandwidth requirement between core edge nodes. This information is supplied by the policy manager, which is a storage of committed SLAs/SLSs.

5.3 SLA/SLS negotiation

Having the capability to establish SLAs *rapidly, accurately and automatically* is a significant contribution to the efficiency of a provider. This becomes even more important when the number of services and customers grows. Having adequate SLA-related mechanisms is therefore considered as a competitive edge by the providers/operators. As there are also dependencies between the providers, the SLAs need to be present throughout the set of providers involved, not only towards the end-customer.

Handling QoS and SLA in an efficient manner introduces a number of challenges; a few aspects are illustrated in Figure 20. Several non-technical aspects will also be included in an agreement between the actors. In addition to the data transfer related aspects, issues like customer support and service provisioning will often be covered.

The SLA template is used to capture a set of Service Level Objectives for a service. A Service Level Objective is a representation of the guaranteed level of service offered. It defines an individual objective for example in terms of service metric, threshold values and tolerances. A service metric could be related to the entire service bundle, to a service element or to a single service interface, but is always related to something visible to the customer. Hence, the QoS requirements described for voice, video and data earlier come into play.

The technical part of an SLA is called a Service Level Specification, SLS, although the actual relations between SLAs and SLSs can be more involved. A service can be said to be provided to a customer by a provider. Prior to the service delivery, negotiation would commonly take place.

The components of an SLS can be grouped as (note that this assumes DiffServ-based services):

- Common unit: Describes the terms of offering the service, e.g. identifying the provider, customer, service type, etc. The period of validity is a central component.

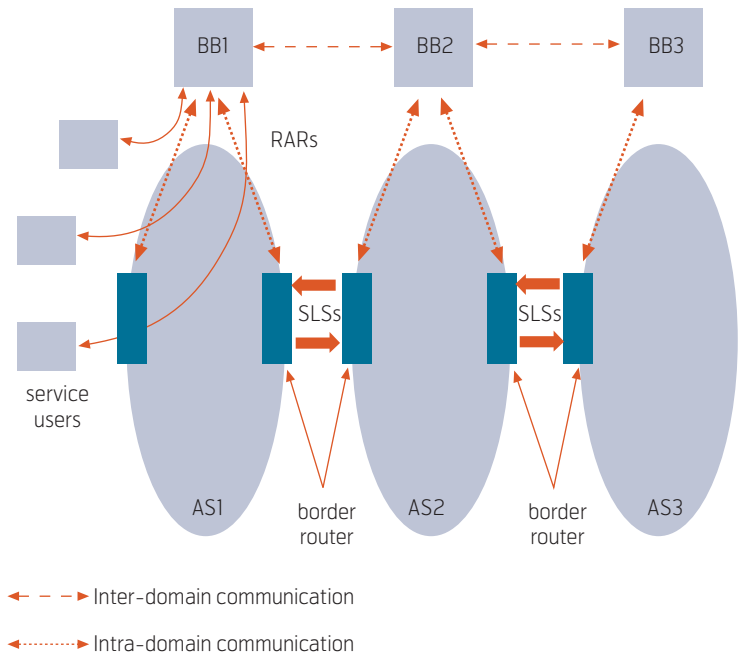


Figure 19 Communication among Bandwidth Brokers

- Topology unit: Describes the nature and number of end-points, further divided into one Service Access Point, SAP, sub-unit and a number of graph sub-units. The SAP sub-unit gives a list of end-points that specify the topology (like hose, pipe or funnel). The end-points can for instance be given by IP addresses. The graph sub-unit gives a list of sources and destinations and how these are related. Unidirectional and bidirectional relations may be described.
- QoS related unit: Describes the traffic flows and the service differentiation provided. Quantitative and qualitative service levels may be given for some or all parts of the topology unit. This unit may further be divided into: i) scope – giving the topology unit (graph sub-unit or end-point) relevant, ii) traffic descriptor – describing the traffic flows (including

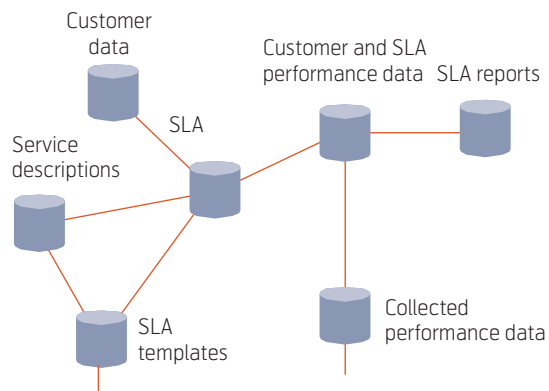


Figure 20 Some of the relevant data for managing SLAs and QoS

DiffServ class, port numbers, protocol information and specification of lower layer), iii) load descriptor – giving the quantity of offered traffic, e.g. given by leaky bucket parameters, as well as treatment of excess of out-of-profile traffic, iv) QoS parameters – delay, jitter and loss of traffic flow.

- Monitoring unit: Defines a set of parameters that are to be collected and reported between the customer and provider. The structure might be similar to the QoS-related unit.

6 Concluding remarks

Is introducing IP QoS worth while? QoS-related mechanisms imply a more efficient network and wider spectre of service portfolio. Hence, savings on the cost of network resources as well as potential additional revenues are within the reach by implementing such mechanisms.

There is a trend towards multiple applications on IP networks covering all components of a multimedia application. In the long run, this requires that QoS-related mechanisms are smoothly operated and thus being a motivation for this paper. Considering a multi-provider environment this becomes more challenging. Industry initiatives, e.g. IPSphere, have been started to address these issues. Hence, reflecting on the motivation and effort, it seems to be more a question of *how fast the concerned IP QoS-related mechanisms need to be introduced.*

References

[22.105] ETSI. *Universal Mobile Telecommunications Systems (UMTS); Service aspects; Services and Service Capabilities*. 2000. (ETSI TS 122 105)

[E.721] ITU. *Network grade of service parameters and target values for circuit-switched services in the evolving ISDN*. Geneva, 1999. (ITU-T Recommendation E.721)

[E.860] ITU. *Framework for a service level agreement*. Geneva, 2002. (ITU-T Recommendation E.860)

[ES.282] ETSI. *Telecommunications and Internet converged Services and Protocols for Advanced*

Networks (TISPAN); NGN Functional Architecture Release 1. Draft, April 2005. (ETSI ES 282 001)

[EU.P806] *EQoS – A Common Framework for QoS/Network Performance in a multi-Provider Environment*. EURESCOM project 806. URL: <http://www.eurescom.de/public/projects/P800-series/P806/default.asp>

[G.1010] ITU. *End-user multimedia QoS categories*. Geneva, 2001. (ITU-T Recommendation G.1010)

[Jens03] Jensen, T. Planning Dependable Network for IP/MPLS over Optics. *Teletronikk*, 99 (3/4), 128–162, 2003.

[Kell00] Kelly, F, Key, P, Zachary, S. Distributed Admission Control. *IEEE Journal on Selected Areas in Communications*, 18 (12), 2617–2628, 2000.

[RFC2475] IETF. *An Architecture for Differentiated Services*. 1998. (RFC 2475)

[RFC2722] IETF. *Traffic Flow Measurement: Architecture*. 1999. (RFC 2722)

[RFC2753] IETF. *A Framework for Policy-based Admission Control*. 2000. (RFC 2753)

[Ross06] Rossebø, J, Sijben, P. Security issues in VoIP. *Teletronikk*, 102 (1), 130–145, 2006. (This issue)

[Ulse06] Ulseth, T, Stafnes, F. VoIP speech quality – better than PSTN? *Teletronikk*, 102 (1), 119–129, 2006. (This issue)

[Y.1540] ITU. *Internet protocol data communication service – IP packet transfer and availability performance parameters*. 1999. (ITU-T Recommendation Y.1540)

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

[Øste2005] Østerbø, O. Performance Modeling of an Upstream Link in Open Access Networks. *Broadband Wireless Access Network Workshop*, Florida, USA, June 2005. (www.ist-oban.org)

Dr. Terje Jensen is Senior Research Scientist at Telenor Research and Development. In recent years he has mostly been engaged in strategy studies addressing the overall network and system portfolio of an operator. Besides these activities he has been involved in internal and international projects in various areas, including network planning, performance modelling/analyses and dimensioning.

email: terje.jensen1@telenor.com

VoIP speech quality – Better than PSTN?

TROND ULSETH AND FINN STAFSNES



Trond Ulseth is Senior Research Scientist at Telenor R&D

As VoIP moves from being an interesting (and cheap) application for enthusiasts to a public service for everybody, the speech quality requirements will be of increasing importance. There are a number of factors that contribute to the user perceived speech quality. Voice over a packet network may introduce new degradations such as packet loss, and increase other degradations such as delay. On the other hand VoIP simplifies the use of wideband (7 kHz bandwidth) codecs offering improved speech quality compared to narrowband (3.1 kHz bandwidth) codecs. Methods to determine the user perceived speech quality are presented, and factors influencing the speech quality of a VoIP connection are discussed. Finally, it is concluded that assuming identical speech coding algorithms VoIP introduces degradations compared to PSTN/ISDN. However, the possibility to implement wideband codecs may give an opportunity for improved speech quality compared to PSTN/ISDN.



Finn Stafsnæs is Research Scientist at Telenor R&D

Introduction

The customers of a VoIP service have expectations to the speech quality. These expectations are based on the experiences made when making telephone calls over the present circuit-switched network (PSTN/ISDN). However, the success of mobile systems offering a lower quality than fixed network systems shows that the users may accept quality degradations if there are other benefits (e.g. mobility). On the other hand, the introduction of the GSM Enhanced Full Rate Codec indicates that there is a demand for a quality that at least is similar to that of the fixed network.

In the beginning VoIP had a reputation for being cheap and of low speech quality. The low speech quality may be acceptable to enthusiasts in the initial phase, but is hardly acceptable on a longer term.

This article presents an overview of how speech quality is measured or assessed and factors that are influencing the user perceived speech quality. Finally, an attempt is made to answer the basic question 'Is the VoIP speech quality better (or poorer) than the PSTN speech quality?'.

Norwegian	Rating	English
Svært God	5	Excellent
God	4	Good
Middels	3	Fair
Dårlig	2	Poor
Svært Dårlig	1	Bad

Table 1 A five point MOS scale with Norwegian and English descriptions

How to measure speech quality?

The term *speech quality* usually means the subjective quality of a conversation. Speech quality can be measured using subjective tests where test persons rate the quality of a communication link. The subjective tests can be

- Conversational tests where two test persons communicate with each other;
- Listening tests where test persons listen to speech signals processed by the system under test.

Listening tests are simpler to carry out, but there are aspects that cannot be tested by using listening tests, e.g. the effects of delay.

A five-point scale is often used to rate the speech quality. A commonly used description for this scale can be found in Table 1. Some tests are carried out using a slide (e.g. on a PC) that gives a continuous scale. However, the quality description used is not changed.

To obtain an acceptable accuracy, a number of test persons have to participate in a subjective test program. The average of all test persons' results is calculated. Statistical analysis is also carried out in order to estimate the accuracy of the results.

Another approach is to compare the original (reference) signal and the processed signal. This approach is called Degradation Category Rating (DCR), while the method described above is called Absolute Category Rating (ACR). An example description for a DCR MOS scale rating is presented in Table 2.

Subjective tests are resource demanding. A lot of effort has been spent on developing models that

Description	Rating
Inaudible	5
Audible but not annoying	4
Slightly annoying	3
Annoying	2
Very annoying	1

Table 2 DCR MOS scale description

makes it possible to calculate the speech quality based on measurements of objective parameters.

For telephony CCITT (now ITU-T) developed an algorithm for calculating the loudness rating for telephone sets; ITU-T Recommendation P.79 [1]. This recommendation is important for the telephony transmission planning in analogue networks. It is also used to specify the sensitivity of telephone sets used in an all-digital network (ISDN). The sensitivity requirements to an ISDN telephone set could be used to specify the sensitivity of a VoIP terminal.

Another standard that describes a method for predicting the subjective quality of narrow-band handset telephony and narrow-band speech codecs is ITU-T Recommendation P.862 [2], which specifies the PESQ (Perceptual Evaluation of Speech Quality) algorithm. PESQ compares the degraded speech with the reference speech and computes an objective MOS value in a 5-point scale.

Both these calculation algorithms address specific parts of a voice connection and do not include effects such as end-to-end delay. An algorithm that addresses the user perceived quality of an end-to-end connection, including end-to-end delay, is the E-model [3]. The calculation algorithm was developed by ETSI

R factor	User satisfaction	ACR MOS rating (lower limit)
90 to 100	Very satisfied	4.34
80 to 90	Satisfied	4.03
70 to 80	Some users dissatisfied	3.60
60 to 70	Many users dissatisfied	3.10
50 to 60	Nearly all users dissatisfied	2.58
0 to 50	Not recommended	

Table 3 Relations between the R factor, user satisfaction and the ACR MOS rating

[4]. The work was presented to ITU-T SG 12, which adopted it and published it as ITU-T Recommendation G.107 [3]. It has been improved and extended by ITU-T SG12. Degradations related to communication over packet networks (e.g. packet loss) have been added. The model calculates a rating factor (*R* factor),

$$R = R_0 - I_s - I_d - I_{e-eff} + A$$

where

R_0 is related to the signal-to-noise ratio;

I_s is related to impairments such as loudness ratings and quantization;

I_d is related to delay;

I_{e-eff} is related to impairments caused by low bitrate codecs;

A is an advantage factor.

The advantage factor is used to take account for user advantages such as mobility.

The present version of the E-model is restricted to narrowband speech communication. There is currently work on extending the algorithm to include wideband speech communication.

ITU-T Recommendation G.109 [5] describes the relations between the *R* factor, user satisfaction and the ACR MOS rating; see Table 3.

Characteristics that influence the user perception of speech quality

The characteristics that influence user perceived speech quality are illustrated in Figure 1. These are

- Speech coding algorithm
- Delay and jitter
- Packet loss
- Echo
- Terminal characteristics (e.g. jitterbuffer, frequency response, noise).

The network related degradations are

- Packet loss
- Delay
- Jitter.

The effects of these degradations on the user perceived quality depends on the application/terminal. As an example, a good jitter buffer implementation may ensure minimal additional packet loss caused by jitter buffer overflow and at the same time minimizing the delay added by the jitter buffer.

Speech coding algorithms

Speech is an analogue signal. To transport speech over a digital network (e.g. a packet-switched network) the analogue signal needs to be converted to digital form. The Nyquist theorem states that to correctly reproduce a digitised signal, the sampling rate has to be more than twice the highest frequency of the signal to be digitised. Furthermore, the speech dynamic range for good voice reproduction requires an analogue-to-digital converter having 13 or 14 bits. Transmission links both in fixed networks and in wireless networks often have a limited capacity. It is therefore necessary to encode the digitised signal to a format that requires lower transmission capacity by implementing a speech codec.

A list of speech codecs for conversational applications is presented in Table 4. Most of these codecs are based on an 8 kHz sampling rate, which means that the maximum speech bandwidth that can be transmitted is close to 4 kHz. Telephone bandwidth is usually somewhat less, 300 – 3400 Hz. Speech transmitted on such a channel is often characterized as *narrowband* speech.

There is also a group of speech coding algorithms that are based on 16 kHz sampling enabling transmission of speech signals covering the frequency band from 50 – 7000 Hz. These codecs are often characterised as *wideband* codecs, and use a 16 bit A/D converter. The dynamic range of these codecs is therefore larger than the dynamic range of the narrowband codecs. ITU-T has recently published a 14 kHz extension of ITU-T Recommendation G.722.1 [6].

In addition to codecs dedicated for speech there are algorithms standardised by other standards organisation such as ISO (International Standards Organization) that covers a wider frequency band than 7 kHz. These algorithms are designed for audio signals that include music, and are outside the scope of this article. Speech codecs are often divided into three classes:

- waveform codecs
- source codecs
- hybrid codecs.

Waveform codecs attempt, without using any knowledge of how the signal to be coded was generated, to produce a reconstructed signal whose waveform is as close as possible to the original. This means that in theory they should be signal independent and work well with non-speech signals. An example is the codec standardised in ITU-T Recommendation G.711 [7], presently the most widely used codec in PSTN/ISDN. This codec uses a non-linear quantizer, giving a 40 % reduction in bitrate compared to a linear quantizer.

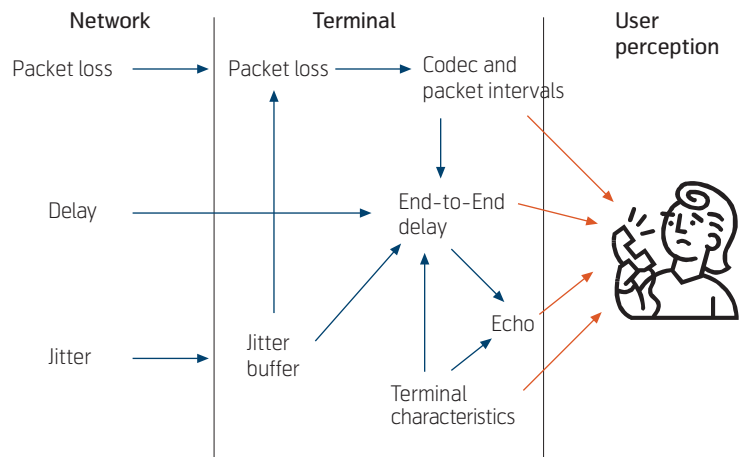


Figure 1 Characteristics influencing user perceived voice quality

Another technique to reduce the required bit rate is to transmit the difference compared with the previous sample instead of the actual sample. This technique is called delta modulation or differential PCM (DPCM). This technique may be further enhanced by predicting the value of the next sample from the previous samples and transmit the difference between the predicted value and the actual sampled value (ADPCM).

The input speech signal may also be split into a number of frequency bands, or sub-bands, and each is coded independently. This is called Sub-band coding. An example is the codec defined in ITU-T Recommendation G.722 [8] where the 7 kHz frequency band is divided into two sub-bands, which are coded independent of each other.

Source codecs operate using a model of how the source was generated, and attempt to extract, from the signal being coded, the parameters of the model. An example is Linear Predictive Coding (LPC). Coders using this technique require very low bitrate, but the quality is usually not good enough for public telecommunication applications.

Hybrid codecs attempt to fill the gap between waveform and source codecs. Although other forms of hybrid codecs exist, the most successful and commonly used are time domain Analysis-by-Synthesis (AbS) codecs. Such coders use the same linear prediction filter model of the vocal tract as found in LPC vocoders. However, instead of applying a simple two-state, voiced/unvoiced model to find the necessary input to this filter, the excitation signal is chosen by attempting to match the reconstructed speech waveform as closely as possible to the original speech waveform. Examples are Multi-Pulse Excited (MPE) codecs and Code-Excited Linear Predictive (CELP) codecs.

Name	Bit rate (kbit/s)	Sampling rate (kHz)	Frame size ¹⁾ (ms)	Remarks
Narrowband codecs				
ITU-T G.711 [7]	64	8	Sample (0.125)	Two companding characteristics, A-law (Europe) and μ -law (North America and Japan) PLC look ahead: 3.75 ms
ITU-T G.723.1 [14]	5.3/6.3	8	30	High rate option: Multipulse Maximum Likelihood Quantization (MP-MLQ) Low rate coder option: Algebraic-Code-Excited Linear-Prediction (ACELP) Look ahead: 7.5 ms
ITU-T G.726 [15]	16, 24, 32 and 40 kbit/s	8	Sample (0.125)	ADPCM
ITU-T G.728 [16]	16	8	0.625	Low Delay CELP
ITU-T G.729 [11]	8	8	10	ACELP Look ahead: 5 ms
3GPP/ETSI AMR [17]	variable	8	20	ACELP Supports: 4.75, 5.15, 5.9, 6.7, 7.4, 7.95, 10.2 and 12.2 kbit/s The 12.2 kbit/s option is equivalent to the GSM EFR (Enhanced Full Rate) codec
IETF RFC 3951 [10] (iLBC)	13.3/15.2	8	30/20	A Global IP Sound (GIPS) developed codec using a block-independent linear-predictive coding (LPC). Robust to packet loss
GIPS G.711 Enhanced [9]	Variable, average 64 kbit/s or less	8	Not specified (sample?)	Proprietary solution. Robust to packet loss
Wideband codecs				
ITU-T G.722 [8]	48, 56 and 64	16	Sample (0.0625)	Two sub-bands are coded independently using ADPCM
ITU-T G.722.1 [6]	24 and 32	16	20	Modulated Lapped Transform (MLT) Look ahead: 20 ms Annex C specifies a 14 kHz mode (32 kHz sampling frequency) transmitting at 24, 32 and 48 kbit/s
3GPP/ETSI AMR-WB [18] (ITU-T G.722.2)	Variable	16	20	ACELP Supports: 6.6, 8.85, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05, 23.85 kbit/s
GIPS iPCMwb [9]	Variable, average 80 kbit/s	16	Not specified (sample?)	Proprietary solution. Robust to packet loss. Compatible with GIPS G.711 Enhanced
GIPS iSAC [9]	Adaptive and variable 10–32 kbit/s	16	Not specified (sample?)	Proprietary solution. Robust to packet loss

Table 4 Speech coding algorithms

There are some speech codecs developed by Global IP Sound, a Swedish-American company that provides embedded speech processing solutions for real-time communications on packet networks [9]. Among their products are four speech codecs that are more robust to packet loss than the codecs standardized by

ITU-T and 3GPP/ETSI. One of these, iLBC, is standardized by IETF [10]. The product portfolio of the company also includes wideband codecs. One of these, iPCM-wb, may interwork with one of the narrowband codecs offered, Enhanced G.711²⁾. It is thus possible to set up a connection between a terminal

¹⁾ Frame size is the speech time interval the codec uses in the encoding process. In addition to this interval there might be a look ahead interval that is used to improve the robustness against packet loss (in conjunction with Packet Loss Concealment).

²⁾ The term Enhanced G.711 is misleading. The codec cannot interwork with G.711 codecs.

using a wideband codec and a terminal using a narrowband codec.

ITU-T has recently begun work on a wideband extension to ITU-T Recommendation G.729 [11] allowing interworking between the narrowband and wideband modes.

Another interesting possibility is to generate wideband speech based on the information received from a narrowband codec. The first ideas have been presented [12], [13], but it is likely that a lot of work is required before a commercial product can be available.

The speech coding degrades the user perceived speech quality. The amount of degradation is a function of the compression ratio, but also on the coding principles used and other implementation aspects. Figure 2 illustrates the MOS rating of some narrowband speech coders [19], [20], [21].

Tandeming of speech codecs (decoding to linear PCM and encoding again) may cause degradation. Examples are illustrated in Figure 3 where tandeming of the AMR codec (at 12.2 kbit/s) causes no significant user perceived quality reduction, while tandeming of two G.729 codecs reduces the MOS rating from 3.9 to 3.3.

Another aspect in an environment where different codecs are used is transcoding. Transcoding usually causes a significant reduction and should be avoided. An example is presented in Figure 3 where transcoding from GSM EFR (equivalent to the AMR codec at 12.2 kbit/s) to G.729 reduces the MOS rating to 3.5, while the ratings of the individual codecs are 4.0 (GSM EFR) and 3.9 (G.729).

There is also a potential for speech quality improvement when using the wideband coding algorithms. Investigations reported by Raake [22] indicate that the wideband bandwidth extension quality improvement could be around 1.0 on the MOS scale. The actual speech codec could reduce this value.

Delay and jitter

The delay sources of a multimedia connection on an IP network are

- Transmitting terminal delay. The main sources are the speech processing (codec) and the speech packetization;
- Network delay;
- Receiving terminal delay. The main source is the receive jitter buffer.

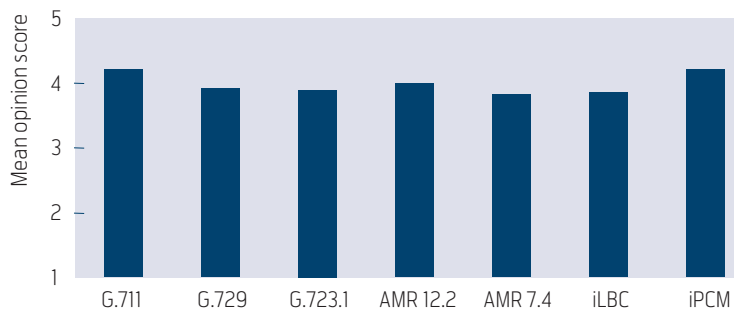


Figure 2 Narrowband speech coders MOS rating

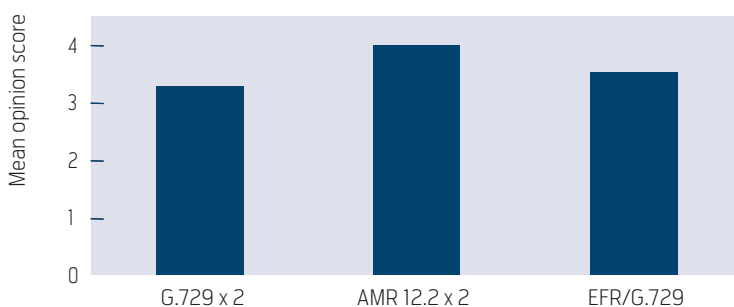


Figure 3 Narrowband speech coders tandeming MOS rating

The speech coding related delay depends on whether the coding algorithm is sample-based or frame-based. The *sample-based algorithms* are low-delay algorithms, introducing less than 10 ms delay (usually 3 ms or less). The *frame-based algorithms* segment the speech signals into frames that typically are 20 ms long; however, there are standardised algorithms that use 10 ms or 30 ms frames.

To reduce the effect of packet loss, Forward Error Control / Packet Loss Concealment can be used. To do so some codecs include an extra time window called look-ahead. The minimum delay introduced by a frame-based algorithm is

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

The duration of a speech packet is flexible. RFC 3551 [23], the IETF Standard that defines the profiles for the Real Time Protocol (RTP) defined in RFC 3550 [24], recommends a packet duration of 20 ms except for the ITU-T Recommendation G.723.1 [14] codec. For G.723.1 30 ms is recommended. This is because the packet duration has to be a multiple of the coding algorithm frame size, and the frame size of the G.723.1 codec is 30 ms. However, the mentioned values are recommended values, implementers may choose the value that is best adapted to the applica-

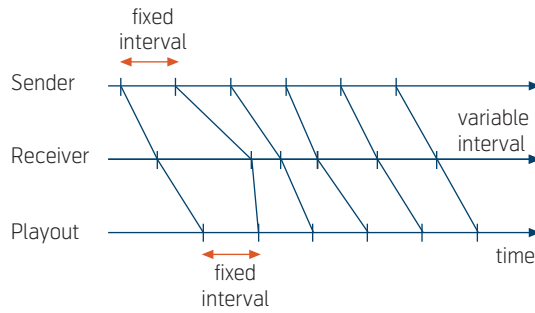


Figure 4 Jitter buffer operation

tion. As an example 60 ms packet intervals are often used in videophone applications because the video coding delay is larger than the audio coding delay.

If multiple speech frames belonging to a frame-based algorithm are grouped together into a single packet, the extra delay will be the duration of one speech frame for each additional speech frame added to the packet:

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

where N is the number of speech frames in each packet.

The core network delay sources are the delay caused at each router of the network connection and the propagation delay. TIA/EIA-TSB116 [25] indicates that the router related delay is approximately 1.5 ms per hop.

The propagation delay depends on the technology used. Table A.1/G.114 of ITU-T Recommendation G.114 [26] presents planning values for calculating propagation delay for various transmission technologies.

The access network may be a significant delay contributor, subject to the technology used. As an example the delay introduced by ADSL may be more than 10 ms depending on the ADSL link capacity. For asymmetrical systems the delay upstream is larger than the delay downstream. For other access technologies it is likely that the delay is less for comparable bitrates.

As stated in the introduction of this section, the main delay source at the receiving terminal is the *jitter* buffer. Jitter is defined as the delay variation caused by queuing in network elements or by routing the packets along different network paths. The speech packets that are sent from the transmitting terminal

at constant intervals are received at variable intervals. Packets might even be out of order. This delay variation needs to be removed and packets need to be reordered before replaying the audible signal to the human user. This is achieved by inserting a buffer (jitter buffer or playout buffer) at the receiving terminal. The function of the jitter buffer is illustrated in Figure 4.

The jitter buffer size needs to match the amount of jitter at the receiving terminal. When the jitter buffer is too short, packets may arrive too late and will be lost. On the other hand, a long jitter buffer increases the end-to-end delay perceived by the user.

The jitter buffer size can be fixed or adaptive. In most scenarios an adaptive jitter buffer is preferable because the jitter characteristics may depend on the actual connection and traffic scenario. However, there are challenges related to adaptive jitter buffers. It is important that the algorithm detects rapid changes in the jitter. In most implementations the adjustment of the jitter buffer size takes place during pauses in the speech. The consequence is that the jitter buffer size is fixed for each talkspurt. To overcome this problem there has been a lot of effort on optimizing the adaptive mechanisms.

A rule of thumb proposes that the jitter buffer should at least be twice the packet intervals. When an adaptive jitter buffer is implemented this value could be the starting point for an algorithm that reduces or increases the jitter buffer size according to the amount of jitter of the incoming packet stream. However, experiments carried out by Liang et al. [27] show that it is possible to adjust the playout of each individual packet by scaling (compressing/expanding) the packets. The packets can be scaled from 50 % to 200 % of their original size without degrading sound quality.

The speech coding algorithm packet loss robustness may also influence the design and size of the jitter buffer. The average delay caused by jitter can thus be reduced compared to traditional jitter buffer solutions. GlobalIPSound [9] offers solutions that probably are based on these principles.

The effects of delay on interactive two-way speech communication are addressed in ITU-T Recommendation G.114 [26]. The Recommendation states that it is desirable to keep the delays seen by user applications as low as possible. Although a few applications may be slightly affected by end-to-end delays of less than 150 ms, if delays can be kept below this figure and there is sufficient echo attenuation, most applications, both speech and non-speech, will experience

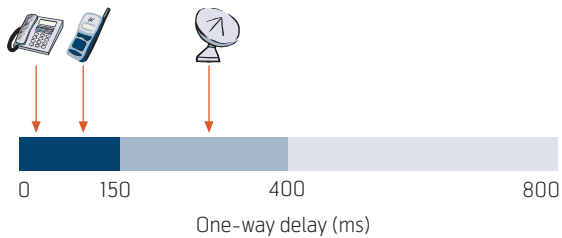


Figure 5 One-way delay examples

essentially transparent interactivity. The upper delay limit for planning purposes is 400 ms. It is however recognised that in some exceptional cases (e.g. double satellite hops) this limit will be exceeded.

Typical speech communication delay examples are illustrated in Figure 5. The delay of a voice connection on a fixed circuit-switched network within Norway is 25 ms or less. The delay of a connection between a terminal in the fixed circuit-switched network and a GSM terminal is approximately 100 ms. The delay introduced by a connection via a geostationary satellite is approximately 270 ms.

A delay estimate of typical VoIP connections within Norway could be as shown in Table 5.

Packet loss

Packet loss degrades the perceived speech quality. The amount of degradation depends on the robustness of the speech codec, and whether or not protection mechanisms such as Packet Loss Concealment (PLC) are implemented. Test results presented to ETSI [28] illustrate both the degradation caused by packet loss and the effects of PLC. Figure 6 describes the effects of packet loss on MOS (Mean Opinion Score) for some relevant codecs with and without PLC.

The tests described above are made when single packets are lost with random distribution. In real networks bursts of packets are quite frequently lost, not single packets, due to effects such as network overload, router queuing and radio transmission disturbances. Subject to the size of the burst (number of consecutive packets lost), burst packet loss may degrade the speech quality more than single packet loss. An article by Clark [29] presents a review of the effect of burst packet loss. The effect is illustrated in Figure 7. The codec used and the burst generating method are not identified in the article.

Jiang and Schulzrinne [30] compare packet loss repair methods and effect on perceived speech quality

	G.711 (ms)	G.723.1 (ms)
Coding and packetization	20	67.5
Processing	1	ca. 10
Access network (ADSL)	15	15
Core network	20	20
Jitter buffer	40	60
Processing	1	ca. 10
Sum	95	180

Table 5 Delay estimate of a VoIP connection within Norway using G.711 or G.723.1 codecs

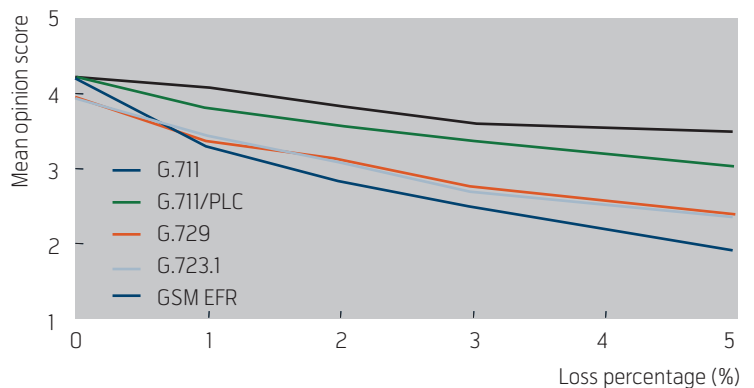


Figure 6 Effects of packet loss on speech quality (MOS)³⁾

under bursty loss. Both calculation results using the E-model and results from subjective tests are reported. Most of the tests were carried out using the ITU-T Recommendation G.729 [11] speech coding algorithm.

The basic loss pattern is specified at 20 ms packet interval. To generate burst packet loss the Gilbert model is used. There are indications that this model is too simple, in ETSI TS 101 329-5 v1.1.1 a four-state Markov model is suggested. (The Gilbert model is a two-state model.)

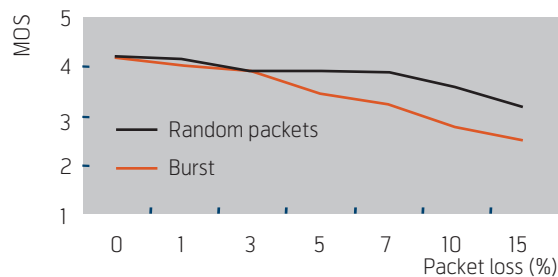


Figure 7 Effects of burst packet loss on MOS

3) Packet intervals are 20 ms for all codecs except for G.723.1. The packet intervals for this codec are 30 ms.

To correctly simulate the loss pattern at larger packet intervals, every second event is picked to simulate 40 ms packet intervals, every third event to simulate 60 ms packet intervals and so on. The consequence of this approach is decreased burstiness when the packet interval increases.

Two packet loss repair methods were compared;

- Forward Error Correction (FEC);
- Low Bit-rate Redundancy (LBR), where a redundant but lower quality version of the same signal is transmitted in subsequent packets.

When no packet loss repair mechanism is implemented the test results show that the difference between the random loss rating and the bursty loss rating is between 0.2 and 0.4 on a five-point MOS rating. It is also found that when carrying out listening tests 40 ms packet interval is rated better than 20 ms packet interval. When carrying out an analysis of delay effects using the E-model, the conclusion is that 20 ms and 40 ms packet intervals could be considered equal, while the performance when using longer packet intervals is poorer.

Global IP Sound [9] offer solutions that are claimed to perform better than the mechanisms described above. The product portfolio includes both speech coding algorithms that are robust to packet loss, and a generic robustness enhancement unit that provides increased packet loss robustness for low bit-rate codec. The solution predicts neighbouring packets from the current one at the price of 2.4 kbit/s addition in bit rate. As long as there is no packet loss, there is no

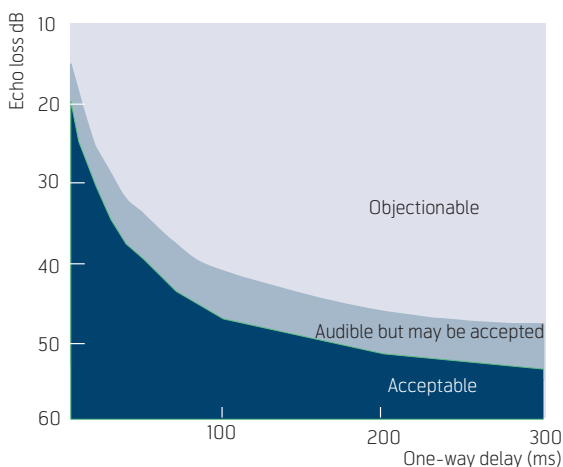


Figure 8 Echo tolerance as a function of one-way delay

extra delay. In packet loss scenarios there will be an extra delay that at maximum is equal to the frame size.

The robustness enhancement unit supports both Global IP Sound codecs and ITU-T standardised codecs such as G.723.1 [14] and G.729 [11].

Echo

Echo is the original speech signal reflected back to the source. The user prefers to hear their own voice; it is an indication that the telephone is working OK. This effect is called sidetone. The user prefers a sidetone that is attenuated 10–15 dB compared to the input signal. The sidetone is not delayed (delay less than 5 ms).

In a telephone conversation a talker can sometimes hear his own voice as a delayed echo. This phenomenon is referred to as talker echo. The effects of an echo depend on delay and the strength of the reflected signals.

In traditional analogue telephony the main talker echo source is the 2/4 wire hybrid. An end-to-end telephone connection over a digital network such as ISDN or IP network is 4 wire. The non-existence of a 2/4 hybrid does not mean that there is no echo problem; the echo is generated in the terminal⁴⁾. There are two sources;

- the acoustic coupling from the receiver (loud-speaker) to the microphone of the terminal;
- the electrical coupling between the wires of the handset cord.

Tests carried out when ISDN telephones were introduced about 15 years ago showed large echo attenuation variations between different handset telephones. These tests showed that it is possible to design telephone handsets where the echo attenuation is acceptable without any echo control mechanism implemented for low delay ISDN connections, but short handsets may require an echo control mechanism.

Headsets are often used with softphone (a PC or PDA with VoIP software). The echo generated by a headset is normally lower than the echo generated by a handset. When using headsets echo cancellers may therefore not be required. Hands-free telephones always require an echo control mechanism.

ITU-T Recommendation G.131 [31] provides guidance on the effect of talker echo. Curves describing

⁴⁾ Analogue telephone adapters have a 4/2 wire hybrid. The main echo source when using these adapters is the hybrid, not the telephone connected to the adapter.

the relation between the required Talker Echo Loss Rating (TELR) and mean one-way delay of the connection are given in the Recommendation. Figure 8 describes the limits for acceptable echo and objectionable echo as a function of one-way delay.

Figure 8 shows that the annoyance caused by echo increases when the connection delay increases. The TELR of a VoIP terminal therefore has to be larger than the TELR of an ISDN terminal. It is therefore likely that acoustic echo cancellers need to be implemented.

Echo canceller technology is used in mobile handsets, and can be seen as a mature technology.

Is the VoIP speech quality better than the PSTN speech quality?

There are VoIP enthusiasts claiming that VoIP speech quality is better than the speech quality of a PSTN connection. The statement might be true – and it might be wrong.

First of all, let us clarify the term PSTN (Public Switched Telephone Network). This term is usually associated with the *analogue telephone network* where the connection between the end user terminal (telephone) and the network is analogue. In most networks today the speech is digitised at the local exchange and transported in digital format (PCM encoded as specified in ITU-T Recommendation G.711 [7]) to the local exchange to which the remote user is connected. The digital signal is then converted to analogue format, and is transported in analogue format to the remote user.

ISDN is in principle a public switched network too, however the communication is digital end-to-end; there is no degradation related to the local loop. The speech coding algorithm is normally PCM as specified in ITU-T Recommendation G.711 [7].

Both PSTN and ISDN telephony are limited to 3.1 kHz bandwidth (i.e. narrowband speech). There is however a possibility to use wideband codecs in ISDN. This option is not frequently used for telephony, but is popular among ISDN videophone/ videoconference users.

In PSTN there are degradations of the speech signal at both ends of the communication link in terms of attenuation and attenuation distortion⁵⁾. Analogue terminals compensate for average local loop loss. The terminal speech signal sensitivity may therefore be

too high or too low, depending on the actual characteristics of the local loop. There is no such degradation in ISDN, it is digital end-to-end.

As described in this article, the main VoIP degradation sources relative to telephony in circuit-switched networks are delay and packet loss. In public networks meeting the objectives of ITU-T Recommendation Y.1541 [32], and assuming proper repair mechanisms are implemented, packet loss does not contribute to significant speech degradation. Using the E-model to estimate the quality degradation caused by pure delay, an increase from 20 ms to 95 ms end-to-end delay corresponds to a MOS reduction between 0.1 and 0.2 when the echo loss is 55 dB or more. This echo loss usually requires implementation of a good acoustic echo canceller. When the echo loss is 46 dB, a loss that often has been used as a target for ISDN terminals, the MOS reduction at 95 ms is approximately 0.7.

Figure 9 presents MOS estimates as a function of delay for three echo loss conditions. MOS = 3.6 is often used as a limit for minimum acceptable quality of a public telephone service.

The user perceived quality of the narrowband speech coding algorithms that are relevant for VoIP is at best similar to the quality of ISDN when the IP connection meets the requirements of ITU-T Recommendation Y.1541 [32]. This comparison is based on subjective tests carried out without packet loss. The packet loss robustness varies, but packet loss will always cause some speech quality degradation.

In an all-IP voice connection it is straightforward to use wideband (7 kHz) speech coders offering a quality that is better than narrowband PSTN/ISDN. However, the speech quality will not be better than wide-

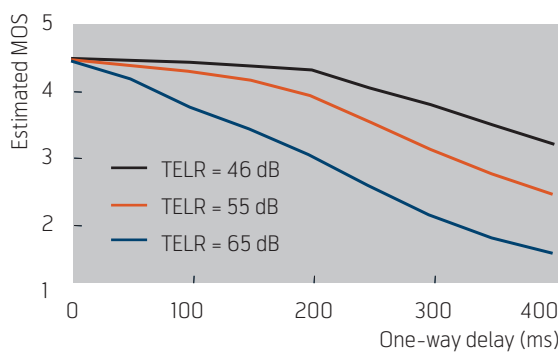


Figure 9 Estimated MOS as a function of echo loss and delay

⁵⁾ The attenuation of a cable in the local loop is frequency dependent. The difference in loss compared to a reference frequency is attenuation distortion.

band telephony service in ISDN. For connection to other networks (e.g. circuit switched PSTN) the possibility to use wideband speech may not be available.

It can be concluded that VoIP may offer speech quality equal to or better than the PSTN when all elements of the VoIP are properly designed. The VoIP speech quality may also be almost equal to the ISDN speech quality when similar class of codecs are used, however the end-to-end delay of a VoIP connection is always larger than the end-to-end delay of an ISDN voice connection. However, wideband codecs is an opportunity to improve the user perceived speech quality.

References

- 1 ITU-T. *Calculation of Loudness Ratings for Telephone Sets*. Geneva, 1999 (ITU-T Recommendation P.79)
- 2 ITU-T. *Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs*. Geneva, 2001 (ITU-T Recommendation P.862)
- 3 ITU-T. *The E-model, a computational model for use in transmission planning*. Geneva, 2005 (ITU-T Recommendation G.107).
- 4 ETSI. *Transmission and Multiplexing (TM); Speech communication quality from mouth to ear for 3.1 kHz handset telephony across networks*. Sophia Antipolis, 1996 (ETR 250)
- 5 ITU-T. *Definition of categories of speech transmission quality*. Geneva, 1999 (ITU-T Recommendation G.109)
- 6 ITU-T. *Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss*. Geneva, 2005 (ITU-T Recommendation G.722.1)
- 7 ITU-T. *Pulse code modulation (PCM) of voice frequencies*. Geneva, 1988 (ITU-T Recommendation G.711)
- 8 ITU-T. *7 kHz audio coding within 64 kbit/s*. Geneva, 1988 (ITU-T Recommendation G.722)
- 9 *Global IP Sound*. (2005, June 30) [online] – URL: <http://www.globalipsound.com/>
- 10 IETF. *Internet Low Bit Rate Codec (iLBC)*. 2004 (RFC 3951)
- 11 ITU-T. *Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear-prediction (CS-ACELP)*. Geneva, 1996 (ITU-T Recommendation G.729)
- 12 Jax, P, Vary, P. On the use of artificial bandwidth extension techniques in wideband speech communication. *The 2nd ETSI Workshop on Wideband Speech Quality Assessment and Prediction*, Mainz, 22–23 June 2005.
- 13 Bernd, I. Bandwidth extension of telephone band-limited speech signals. *The 2nd ETSI Workshop on Wideband Speech Quality Assessment and Prediction*, Mainz, 22–23 June 2005.
- 14 ITU-T. *Dual rate speech coder for multimedia communication transmitting at 5.3 and 6.3 kbit/s*. Geneva, 1998 (ITU-T Recommendation G.723.1)
- 15 ITU-T. *40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM)*. Geneva, 1990 (ITU-T Recommendation G.726)
- 16 ITU-T. *Coding of speech at 16 kbit/s using low-delay code excited linear prediction*. Geneva, 1992 (ITU-T Recommendation G.728)
- 17 3GPP. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Mandatory speech CODEC speech processing functions; AMR speech CODEC; General description (Release 6)*. Sophia Antipolis, 2004 (3GPP TS 26.071v6.0.0)
- 18 3GPP. *3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Speech codec speech processing functions; Adaptive Multi-Rate Wideband (AMR-WB) speech codec; General description (Release 6)*. Sophia Antipolis, 2004 (3GPP TS 26.171v6.0.0)
- 19 ETSI. *Digital cellular telecommunications system (Phase 2+); Performance Characterization of the GSM Adaptive Multi-Rate (AMR) speech codec (GSM 06.75 version 7.2.0 Release 1998)*. Sophia Antipolis, 2000 (ETSI TR 101 714 v7.2.0)
- 20 Perkins, M, Evans, K, Pascal, D, Thorpe, L. Characterizing the subjective performance of ITU-T 8 kb/s speech coding algorithm – ITU-T G.729. *IEEE Communications Magazine*, 35 (9), 1997.
- 21 Campos Neto, S F, Corcoran, F L, Karahisar, A. Performance assessment of tandem connection of enhanced cellular codecs. *1999 IEEE International Conference On Acoustics, Speech, and Sig-*

- nal Processing (ICASSP99)*, Phoenix, Arizona, 15–19 March 1999.
- 22 Raake, A. How Much Better Can Wideband Telephony Be? Estimating the Necessary R-scale Extension. *ETSI Workshop on Wideband Speech Quality in Networks and Terminals Assessment and Prediction*, Mainz, 8–9 June 2004.
 - 23 IETF. *RTP Profile for Audio and Video Conferences with Minimal Control*. 2003 (RFC 3551)
 - 24 IETF. *RTP: A Transport Protocol for Real-Time Applications*. 2003 (RFC 3550)
 - 25 TIA/EIA. *Voice Quality Recommendations for IP telephony*. Arlington, USA, 2001 (TIA/EIA TSB116)
 - 26 ITU-T. *One-way transmission time*. Geneva, 2003 (ITU-T Recommendation G.114)
 - 27 Liang, Y J, Färber, N, Girod, B. Adaptive Playout Scheduling and Loss Concealment for Voice Communication Over IP Networks. *IEEE Transactions on Multimedia*, 5 (4), 1993.
 - 28 ETSI. *Telecommunications and Internet Protocol Harmonisation over Networks (TIPHON); Actual measurements results*. Sophia Antipolis, 2002 (ETSI TR 101 329-6 v2.2.1)
 - 29 Clark, A. Modelling the effects of Burst Packet Loss and the Recency on Subjective Voice Quality. *The 3rd IP Telephony Workshop 2002*, New York, 28 April – 2 May 2002.
 - 30 Jiang, W, Schulzrinne, H. Comparison and Optimization of Packet Loss Repair Methods on VoIP Perceived Quality under Bursty Loss. *NOSS-DAV'02*, Miami Beach, 12–14 May 2002.
 - 31 ITU-T. *Talker echo and its control*. Geneva, 2003 (ITU-T recommendation G.131)
 - 32 ITU-T. *Network performance objectives for IP-based services*. Geneva, 2002 (ITU-T Recommendation Y.1541)

For a presentation of the authors, please turn to page 2.

Security issues in VoIP

JUDITH E.Y. ROSSEBØ AND PAUL SIJBen



Judith E.Y. Rossebø is Research Scientist at Telenor R&D

In the past year, numerous service providers, including the incumbent Telenor, have launched VoIP services in Norway. According to the mainstream media, service providers and users alike stand to save substantially by migrating to VoIP instead of relying on traditional circuit switched telephony (as provided in the GSM and PSTN/ISDN networks), and yet the question remains, what about security? What are the concerns? What are the risks at stake? This article addresses the main security issues and challenges for VoIP. The main security issues for VoIP are authenticity, privacy, and availability. We highlight a number of important threats involved, and countermeasures to these threats that may practically be implemented are discussed. Additionally, implementation considerations regarding the widespread use of Network Address Translators (NATs) and Firewalls in customer networks, are addressed.



Paul Sijben is chief technologist of EemValley Technology

Introduction

VoIP is a technology for producing telephone services on IP-based networks. Traditionally, these telephone services have been provided by the public switched telephone network (PSTN/ISDN), which has been managed and completely controlled by single, national operators in each country in Europe, and for GSM the situation initially was similar. The risks were known, and managed.

Since the mid '90s this situation has been evolving in Europe. The national operators still exist, but in addition second operators, third party vendors, ISPs and mobile providers (GSM, GPRS, UMTS) are also interconnecting and providing a multitude of services such as VoIP, video conferencing, video on demand etc. This dynamic new situation has given rise to new threats and risks, which are more complex and unpredictable than for any one service or technology in isolation. This enriched threat model requires extensive countermeasures in order to be able to deliver services of acceptable quality, reliability and security. In this paper we introduce the security issues inherent in this new complex situation and address how some of them can be mitigated.

VoIP caused a lot of excitement towards the end of the 90s, with the promise of providing a viable technology for the migration from the monolithic public switched telephone network (PSTN/ISDN) to next generation networks, for which telephone services are produced on an IP-based network. At the turn of the millennium, it was announced that the IETF's Session Initiation Protocol (SIP) standard would be chosen as the basis for the 3GPP IP multimedia subsystem (IMS). SIP at this point, was still in an early

phase of development. Problems with poor voice quality for the early Internet-based offerings, along with the added barrier of cumbersome technology, e.g., having to phone from the PC made it difficult for consumers to embrace the new technology, and lead to slow adoption rate. The immaturity of the emerging SIP standard contributed largely to the slow down of the roll out of VoIP services along with uncertainty in the economic and market related factors, and the lack of a solid business model.

Today, VoIP is being used everywhere with different levels of success. Home users may use an Analogue Terminal Adapter (ATA) to use their legacy POTS telephone sets and make telephone calls over the Internet. PC users have a choice of applications that allow them a rich user experience and address book facility, and VoIP telephones are available both as desktop models and cordless handsets using Wi-Fi. Mobile nomadic users may use their VoIP accounts wherever they find a broadband Internet connection.

As is usually the case in software and systems development, VoIP security has not received sufficient attention during the development phases and is lagging behind in the deployment.

VoIP security the next challenge

It is important to realize that VoIP still has a minority of subscribers compared to the PSTN/ISDN and GSM subscribers. The Norwegian Post and Telecommunication Authority (PT) reported in the statistics for the first half of 2005, a total of 106,500 IP telephony subscribers per 30/06/06 [1].¹⁾ While these numbers indicate a rise in popularity of and interest in VoIP, the numbers are still insignificant with respect

¹⁾ However, it should be noted that there is a large grey area as the numbers of e.g. Skype and Skype out users are unknown.

to the number of PSTN/ISDN and GSM mobile telephone subscribers in Norway. Therefore it is too early to draw any definitive conclusions about VoIP security impact at this point.

This article addresses major security issues of VoIP; charging fraud prevention, protection of privacy, and availability. Regulatory requirements present challenges such as providing sufficient protection of the regulated service from abuse via the non-regulated VoIP services, and ensuring the availability of regulated services, making this a security issue. Additionally, the widespread use of NATs and Firewalls in customer networks, although not a direct security issue, but a related one, presents problems for VoIP.

In this article we give an overview of the threats and proposed countermeasures while also addressing key technological issues such as traversing NATs and Firewalls. Regulatory issues are also addressed as this clearly relates to the required level of security in the network. We provide an overview of the state of the art in standardization and international forums, and point out important initiatives to watch in 2005/2006.

The remainder of this paper is structured as follows: We first highlight the security issues surrounding VoIP, some of these are inherent in the telephone service and some are inherent in the use of IP and some are typical to VoIP. After addressing the security state of the art for VoIP we address a number of countermeasures that may be practically employed. We then address implementation issues regarding the related topic of NATs and firewalls, devices that were intended to keep networks secure, but are hampering the deployment of VoIP. The paper is wrapped up with an overview of where VoIP security is addressed in the standards.



Figure 1 Captain Crunch Whistles

VoIP threat model

Telephone fraud and hacking incidents, a brief historical view

Abuse and misuse of the telephone system is not a new phenomenon. In this section we present a set of known threats that are still valid in the VoIP world.

A well-publicized threat is abuse of the service for the purpose of charging fraud. In the 1960s and 1970s it was discovered that Captain Crunch cereal giveaways, small plastic whistles, could be used to gain access to free telephone calls. By blowing the whistle into the receiver, the AT&T systems would allow free calls. This discovery, by amongst others, John "Captain Crunch" Draper [2], led to widespread charging fraud against the company.

During the 1980s, the first so-called hacking incidents began to emerge. In 1988 R.T. Morris introduced the first worm into the Internet, K. Mitnik was arrested for breaking into the Digital Equipment Corporation, and K. Poulsen was arrested on phone tampering charges, involving jamming of phone calls on a large scale. During the 1990s, Mitnik frequently accessed the public telephone network in Las Vegas, and although he managed his hacking feats due to technical prowess, he also applied numerous social engineering techniques to obtain the information he needed to hack his way into the telephone systems.

The first example of an attack on a telecommunications system in the USA with implications for national security was carried out in 1997. A juvenile hacked into the local exchange that serviced the Worcester, Massachusetts area and managed to shut down the PSTN for 600 local users as well as disrupting the fire, police, and 911 emergency services. The investigation that followed revealed that the vulnerability that made the attack possible was present in over 20,000 telephone exchanges across the USA [3].

New threats faced by VoIP

As the Internet and other IP-based networks have expanded and become widely used, the opportunities for abuse have also grown. The threats due to exposure to the Internet are far greater than the threats in the traditional telephony world. This is because geographical distances (and related long-distance rates) are never a deterrent. Like any other IP-based application, VoIP is vulnerable to the same kinds of attacks that are widespread across the Internet today.

The IP infrastructure is also considered to be very brittle. Denial of Service (DoS) attacks against SIP servers and their supporting DNS/ENUM infrastructure are easily mounted and difficult to counter or

mitigate. Broadband IP networks today are not (yet) known for their robustness despite the defence legacy of the Internet.

Primary Goals of VoIP Security

The commercial objectives of any commercial network that impact on security are to ensure profitability of the network, availability of the network and customer confidence. This is to be realized by addressing the following technical issues.

These commercial objectives break down to the following technical security issues for VoIP; charging fraud, protection of privacy, and ensuring availability of the VoIP services. The goals for VoIP should therefore aim to reduce these risks by reducing the ability to mount these attacks and to limit their impact.

We therefore define the following technical objectives for VoIP Security:

- *Prevention of masquerade.* This means being able to determine that a user claiming to be Alice is always Alice, Bob is always Bob, and Bob cannot pretend to be Alice, and this applies to both masquerade of the user and of the system or service.
- *Ensure availability of the VoIP services.* By this we mean that the service must be accessible and usable on demand by an authorised entity. This is crucial for e.g. emergency services. In general, a user expects to be able to place a call and complete the call without being cut off in the middle.
- *Maintain privacy of communication.* In many cases, the parties to a call communicate across

public networks, and mechanisms must be in place to prevent eavesdropping. Furthermore, the only delivery points for communication have to be the legitimate parties to the call.²⁾

Given these objectives, we first examine the threats to VoIP, and then describe countermeasures that can be implemented to reduce the threats to try and meet these objectives.

Regulatory Issues

The PSTN/ISDN public telephony service is regulated. This means that the state regulatory authorities set requirements for both the service functionality and the service quality, and the provider of PSTN/ISDN services must follow those requirements. In general, IP-based services are currently not regulated; however, the situation for IP-based telephony regulation is currently evolving. This is due to the emerging category of IP-based telephony that enables customers to both receive calls from and terminate calls to the public switched telephone network (PSTN) and can therefore be categorized as a public telephone service.

The United States Federal Communications Commission (FCC) voiced in February of 2004 the intent to begin regulating⁴⁾ the IP-based services environment by making the decision to require reporting of major network outages including signalling systems, and on May 19, 2005, the FCC announced the decision to require Interconnect VoIP Providers to provide enhanced 911 services. Note that the requirement at this point does not include VoIP soft-phone applications; however, it is clearly stated that this is a future goal [4].

Recently, the Norwegian Post and Telecommunication Authority (PT) published a report explaining how VoIP applications that are similar to PSTN/ISDN should be regulated [5]. On the basis of this, it is natural to expect that PT will follow a similar path as the FCC.

Regulation means that there are also expectations about the level of security that is required. The connections to the regulated service must of course be protected adequately from abuse via the unregulated service. Ensuring availability of emergency services is crucial. This means that it is important to be able to account for the origin of the call, and the accuracy of

PSTN/ISDN telephony	VoIP
Regulated Service	Not regulated in Norway per May 2005. ³⁾
ITU standards dictate stringent requirements for reliability, availability, performance, and delay	ITU H.323 standard and IETF SIP (RFC 3261) both describing an application protocol but no performance, reliability or QoS.
Closed signalling network (SS7)	Signalling (usually) over the public Internet

Table 1 Some differences between VoIP and PSTN/ISDN

2) Lawful intercept obviously is an explicit exception from this goal.

3) However, this is expected to change and regulation is currently being addressed [5]. For more information, see the article in this issue on VoIP – regulatory aspects from a Norwegian perspective [6].

4) Historically there is a difference in the meaning of regulation between Europe and the USA. In the USA regulation often implies taxation while in Europe it implies customer protection and equal access.

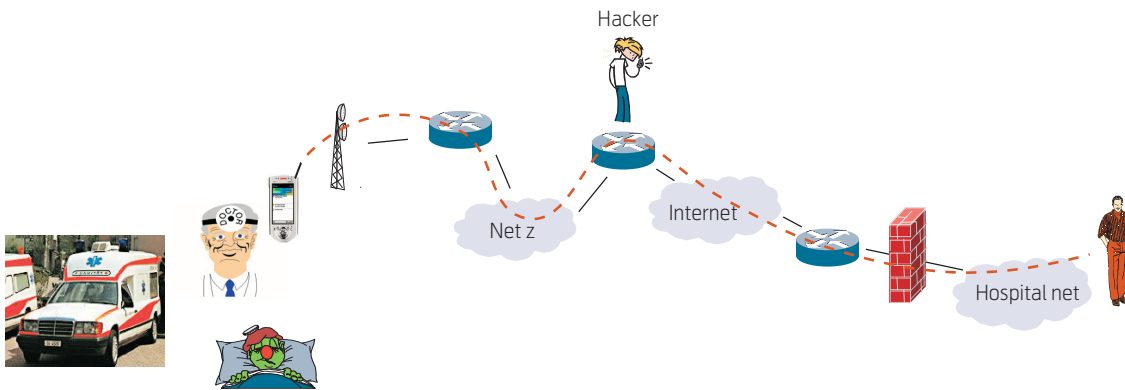


Figure 2 The Emergency service must locate the patient

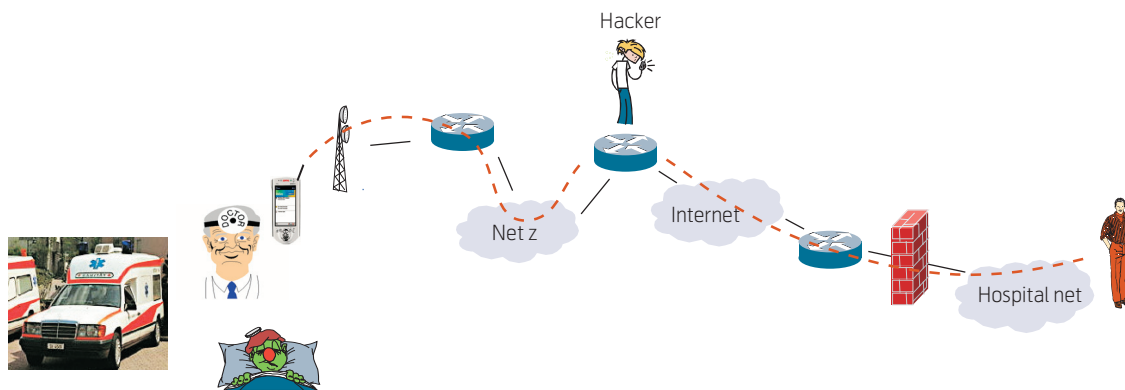


Figure 3 The communication link must be secured against DoS attacks

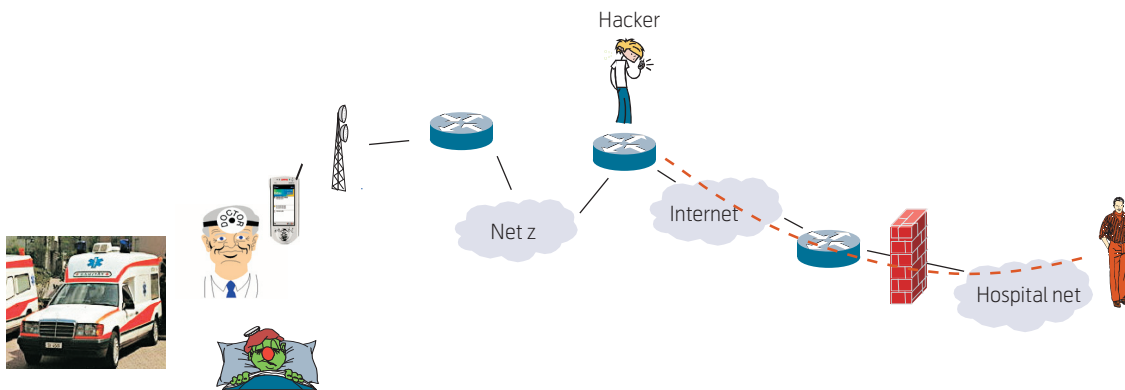


Figure 4 It is important to secure access for the authorized users, and keep out non-authorized users

this information must be ensured. It is crucial that the call is routed correctly to the local emergency service center, and not routed incorrectly to an emergency service center 1000 km away. The communication link must be secured for the duration of the emergency call, and it is important to secure access for the authorised users, and keep out unauthorised users.

Currently, it is possible to deliver a regulatory-compliant IP-based telephony solution that relies heavily on the mechanisms used for securing the PSTN/ISDN, most importantly, the use of the physical line

ID for geographical origin identification. However, a regime must be in place, and an infrastructure must be rolled out to enable meeting the future regulatory requirements for providers of e.g. IP Telephony over WLAN or any VoIP soft-phone applications.

The security threats to VoIP

A threat to VoIP is defined as a potential cause of an unwanted incident, such as telephone fraud or loss of availability of the IP-based telephone service. In this section, we discuss some of the threats to VoIP. Note

that this is not intended to be an exhaustive list, but rather a list of threats that are of importance, and should be addressed by the VoIP service provider.

We have identified the following threat families:

- Masquerading
- Denial of Service
- Eavesdropping
- Abuse of access

Masquerading

A *masquerade* is the pretence of an entity to be another entity. Masquerading can lead to charging fraud, breach of privacy, and breach of integrity. This attack can be carried out by hijacking a link after authentication has been performed, or by eavesdropping and subsequent replaying of authentication information.

Using a masquerade attack, an attacker can gain unauthorised access to VoIP services. An attacker can steal the identity of a real user and obtain access by masquerading as the real user. By employing a replay attack, the attacker can capture the authentication credentials of an authorised user and replay the authentication message at a later time to obtain fraudulent access to a service (and in this case the real user may be charged for the calls placed by the masquerading user). In another form of masquerade, an attacker replaying or masquerading as a service may deceive the user, so that the service the user intended to access is then not available.

The simplest form of masquerade is re-use of username and password that were obtained through interception or social engineering. A more advanced example of how authentication information can be obtained for the purpose of masquerade is by reverse engineering of passwords in the case of SIP digest authentication [7]. The attacker sends false challenges⁵⁾ to the SIP User Agent (UA) in the user's terminal to generate a list that can be used to crack the commonly used Message-Digest algorithm 5 (MD5) [8] cryptographic hash of the password. A Masquerade can then be combined with alteration of data in order to obtain access to services for the purpose of fraud or for placing a malicious call.

Denial of Service

A *denial of service (DoS) attack* is an attack that is conducted to deliberately cause loss of availability of a service. We identify DoS attacks at several levels; transport-level, server level, signalling level.

- Transport level: An IP-level DoS attack may be carried out by flooding a target, e.g. by ping of death or Smurf attack [9].
- Server level: Servers may be rendered unusable by modifying stored information in order to prevent authorised users from accessing the service.
- Signalling level: At the SIP level one may overload the SIP server with too many (possibly invalid) messages, thus making it unavailable to handle legitimate SIP messages. Unauthorised users may also create over-usage problems having an overload effect and this way degrade the quality of service for the authorised users.

An example of an unwanted incident caused by a DoS attack is disruption of network services, e.g. collapse of the entire SIP signalling network. There are also several examples of DoS attacks conducted by misuse of call forwarding services.

Most of the generic IP and Internet-related attacks are well known and we do not discuss them in more detail here. In this section we further elaborate on attacks specific to the SIP protocol. We identify several kinds of attacks that are viable against SIP.

- A DoS attack on a call can be carried out by sending spoofed SIP "bye" messages, to tear down the call, to the UAs participating in the SIP call.
- All of the participating entities (UA, proxy, FW/NAT, Media Gateway) that process SIP signalling are susceptible to DoS attacks by simply flooding the target with "register" or "invite" packets. Simple tests carried out on vulnerability to this type of attack have shown that most SIP implementations are vulnerable to this type of attack.
- SIP is also very susceptible to overloading the server with illegal SIP messages. The free form text-based structure of SIP ensures that any entity aiming to receive a SIP message must investigate the entire SIP message before it may deem it to be valid or invalid as important information related to the validity of the message may appear anywhere in the message.

Eavesdropping

Eavesdropping on signalling or media is a threat whereby an attacker finds a way to copy legitimate messages between the targets. This threat is a threat to privacy; an attacker can gain information in an unauthorised manner such as obtaining personal

⁵⁾ The attacker is able to choose the nonce, making cryptanalysis easier.

information about the origin and destination of the call, overhearing a supposedly private conversation or intercepting personal information such as a telephone-bank customer's account number and PIN (used for conducting bank transactions over the telephone) can be captured and misused.⁶⁾

This attack is easily mounted for e.g. VoIP traffic over WLAN or other shared infrastructures. This is a real threat as there are packet-sniffers readily available with built in codecs which can be used for eavesdropping on VoIP.

As an extension to this threat one may consider manipulation of legitimate communication. Using information gained by eavesdropping on the signalling, an attacker can manipulate fields in the data stream and make VoIP calls for the purpose of fraud or to place malicious calls or inject their own speech to make the other party seem to say certain things they did not intend.

This threat is very likely in VoIP as SIP messages and media streams may be encrypted (i.e. the standards that allow for this exist) but in practice are always sent unencrypted to ensure interoperability or ease of debugging the network.

Abuse of access

Abuse of access is a threat where malicious users/programs abuse their access to the system. Abuse of access can take all kinds of forms. We can identify access to the system by a malicious user and access by a malicious program on interfaces exposed by the system.

Abuse of access is not a threat unique to VoIP. Microsoft Outlook, for instance, exposes an application programming interface (API) that allows other programs to access the user's address book and send emails. When Microsoft introduced this feature along with the ability to run programs attached to emails directly without user intervention, email viruses, unknown until then, were invented overnight.⁷⁾

In VoIP we see this threat re-surface in several forms. One form is abuse of click-to-dial services, where companies will call back users via the regular phone system. This service is usually offered through a webpage where anyone can enter any phone number. This is very much open to abuse and has resulted in many companies discontinuing this option.

Another example of abuse of access is for instance the Skype application. The Skype client application presents an API to other applications allowing them to initiate calls and insert events in existing calls. This API is protected by an access control mechanism where users are asked if they want to allow the application to control the Skype Client. Such access control is known not to be fool-proof, the simplest way to thwart it is to lure users to accept this control similar to luring users to click on certain links in emails. The Skype API is very powerful and allows a third-party application to control all the aspects of the Skype client. Several bugs have already been found in this API, and been subsequently fixed by Skype. The first Skype Trojan has already been found loose on the Internet [10], and experts generally are expecting the first Skype virus to appear soon.

VoIP security – State of the Art

Given these known threats, how does the VoIP service provider determine what the risks are? A risk analysis must be conducted. Critical risks are those that are relatively easy to carry out and which break the system in a way that destroys the viability of the system for the user or the provider. Soft phones are particularly vulnerable to e.g. fraud as the risk that an attacker gains control over a PC/terminal over the Internet is substantial. VoIP can be eavesdropped on if an attacker gains access to the residential network. Softphone clients based on mobile terminals or on personal computers are much more vulnerable to attacks than dedicated hardware telephones or analogue terminal adapter (ATA) configurations. This is because the softphone client is an application running on a computer operating system and is therefore vulnerable to the same attacks as any software application as well as being vulnerable to problems due to other attacks on the PC or mobile terminal such as viruses.

It is up to the VoIP service provider to assess the risks and determine the level of security required for adequate protection.

The current situation for securing VoIP services deployed in fixed networks still relies in part on physical security. For authentication, however, the analogue terminal adapter (ATA) is often authenticated by MAC address or some sort of proprietary code contained in the ATA, and the SIP UA username and password is then passed to the SIP server, for authen-

⁶⁾ In Norway, banks have already made the transition to employing one time password solutions for conducting bank transactions over the telephone, and this commendable advancement is now being adopted elsewhere.

⁷⁾ Note that this threat is not new to Microsoft applications, introduction of a powerful scripting language (that later became Visual Basic) in Microsoft Word around 1990 was the origin of word-processing viruses.

	PSTN/ISDN telephony	VoIP
Authentication	Based on physical line ID, caller ID	Based on logical ID, Username and password ⁸⁾
Confidentiality	Physical security	Encryption techniques
Integrity	Physical security	Encryption techniques
Availability	Physical security/physical access control, including a separate SS7 signalling network	Network/Service access control, DoS protection measures such as Intrusion Detection Systems (IDS)

Table 2 Differences in how security countermeasures are applied

tication as standardised in SIP. The physical address/line Id is used for providing the geographical location for e.g. emergency services.

Although equipment with support for the protection of SIP signalling (SIPS) is available in the Norwegian market, it is generally not in use. The SIP signalling passes in clear text. For ADSL customers connected using an ATA adapter, this means no increase in risk over traditional PSTN/ISDN. However, for calls transported at least in part across WLAN networks, the risk of eavesdropping is real.

The case for securing VoIP in the nomadicity cases is different, in particular there are challenges in the nomadicity cases for emergency services provisioning. The challenge of providing the 112 emergency service in the nomadicity case is routing the call to the nearest emergency service and obtaining the accurate geographic origin of the call. Currently, it is possible to obtain dispensation from the regulation requirements (of e.g. identification of geographical origin) for the nomadicity case.

For VoIP applications deployed in 3G mobile networks, the USIM is used as the basis for securing IP-based multi-media services. This USIM is a removable smartcard chip outwardly similar to the SIM card in today's GSM cellular handsets; however inside, it is very much more advanced. The USIM security functions are exploited for IMS authentication as well as for integrity protection of the IMS SIP signalling.

Considering the above description we can conclude that deployed VoIP security is weaker than the security provided in the traditional PSTN/ISDN telephony world. Table 2 gives an overview. Although traditional POTS handsets are very simple and VoIP handsets are much more complex, intercepting a POTS call or impersonating a POTS caller is much more difficult because one needs to be able to have

physical access to the wire pair that is authorized to make these calls. In the VoIP world one can impersonate a user by logging in from any IP address on the Internet.

Besides the problems with the infrastructure the general problem for VoIP is that user terminals cannot be assumed to be tamper proof and thus login credentials (today typically a username/password combination) may be stolen from the terminal to mount an impersonation attack or to place calls on somebody else's cost. To counter this, such VoIP user terminals should contain a tamper proof hardware token to safely protect user data and authentication and encryption keys as is the case in 2G and 3G based services. However, at this point, the competitive climate for IP-based services indicates that the costs are prohibitive.

Threat analysis

In ETSI STF 292 (see the section in this paper on ETSI TISPAN) we have executed a Threat Vulnerability and Risk Analysis (TVRA) of a VoIP deployment using the session initiation protocol (SIP) and the telephone number resolution system called Enhanced Number (ENUM). The analysis was conducted using the method for systematically analyzing threats, vulnerabilities and risks developed by the STF 292. The method involves a systematic identification of assets and threats and weaknesses, and weighting these to classify the risks.

Target of Evaluation

The target of evaluation (TOE) is a generic standards compliant VoIP deployment, assuming standards-based implementations of SIP and ENUM without deployment specific non-standard security features, as publishing an analysis of a real deployment would be inappropriate. As signalling protocol SIP [11] was chosen and ENUM [12] was chosen for number resolution and routing as many service providers are con-

⁸⁾ Possibility for PKI-based in the future (Verisign, etc.). For 3GPP IMS, VoIP service access is username/password based; however, access to the IMS is based on the 3GPP mechanisms [13].

sidering the combination of these technologies for their deployments.

The deployment architecture is depicted in Figure 5. The figure shows 2 end-user domains from which users are able to make VoIP calls. The end-user terminals either have direct access to the ENUM servers (for instance if they double as the service provider's DNS servers) or just the SIP servers have access to the ENUM servers, the latter case is called Infrastructure ENUM. The SIP servers are reachable for their customers, if these customers need access to the VoIP network while travelling, the servers may be accessible from the whole of the Internet.

TVRA interim results

The final results of this analysis are due to be published by ETSI in 2006. However the interim situation at time of writing is that 50 assets have been identified (both physical assets such as servers and routers and logical assets such as data elements and information in data storage), 15 kinds of threats have been identified, 85 vulnerabilities have been classified, of which 24 lead to critical risks. For this analysis, by critical risks, we mean those for which any connected VoIP user (for a VoIP Internet deployment this means anyone on the Internet) is capable of attacking the VoIP infrastructure such that it impacts many users, causing substantial losses of service availability or enabling massive fraudulent use of services without requiring expertise to do so.

A summary of some of these critical risks found is given in Table 3. The table shows the assets (logical

assets in physical assets), the type of vulnerability the risk applies to, the threat that may impact this vulnerability, the expertise level necessary to execute the threat, the level of access necessary⁹⁾ and the time necessary to mount the attack. This results in a computed likelihood, which combined with the impact¹⁰⁾ when the threat comes true results in the stated unwanted incident.

Each of the threats listed in Table 3 can lead to partial or total loss of service availability. The threats involving manipulation of credentials can also lead to consumer data security breaches. The resulting major unwanted incidents for the service provider are loss of reputation, loss of revenue, and in the case of consumer data security breaches, legal issues.

Examples of successful attack scenarios are; sending malformed messages to overload SIP servers, overloading ENUM/DNS servers by sending many requests, and poisoning ENUM servers into giving the wrong results to other users. Most of these attacks are not new to the Internet. However, by bringing telephony to the Internet one inherits the Internet's threats.

Countermeasures

In this section, we describe countermeasures that can be deployed in order to reduce the possibility for misuse by protecting the vulnerabilities that can be

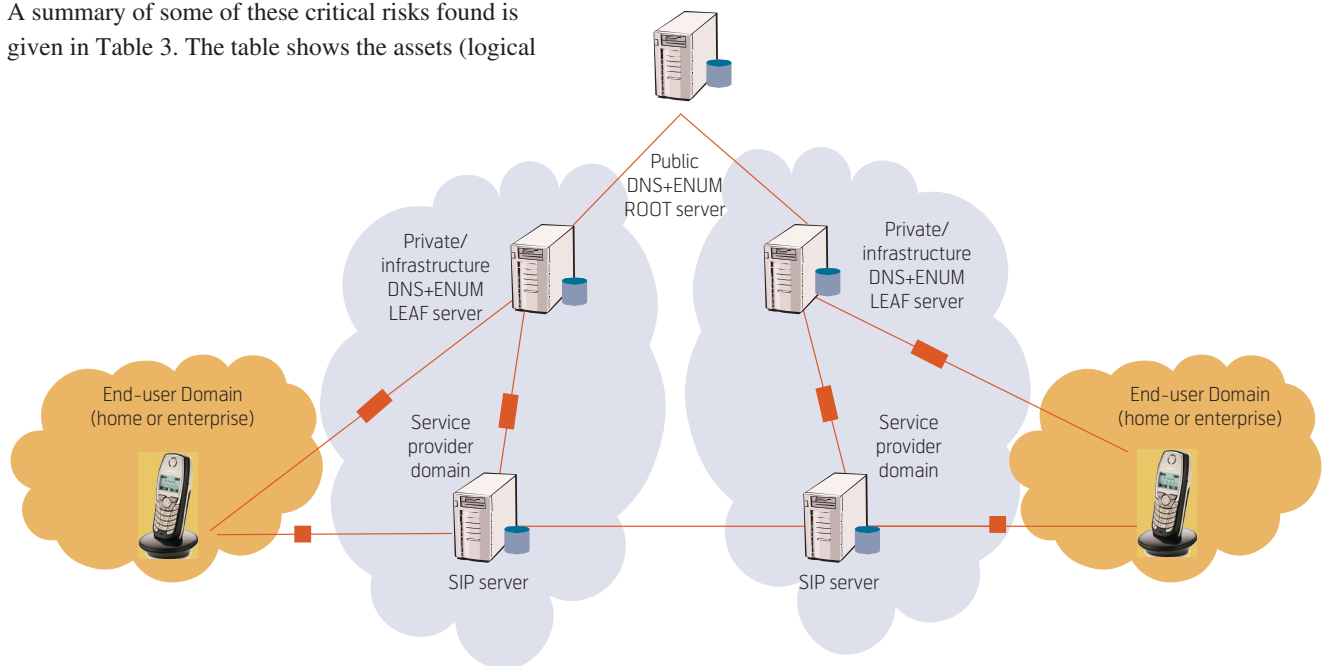


Figure 5 Example VoIP deployment

⁹⁾ Which is unlimited for servers accessible through the Internet.

¹⁰⁾ Impact is high when all or most of the users of the service provider find their service disrupted, low when only one user is affected.

Asset	Vulnerability	Threat Family	Threat	Expertise	Access Description	Time	Likelihood	Impact	Unwanted incidents
Call state IN SIP or other session server	Illegal message content	Closing of sessions	Denial of service	Layman	Unnecessary or unlimited access	<== 1 day	Likely	High	Loss of service availability
Call state IN SIP or other session server	Illegal message format	Overload of communication	Denial of service	Layman	Unnecessary or unlimited access	<== 1 day	Likely	High	Loss of service availability
Data in transit IN link to ENUM leaf server	Limited transport/processing capacity	Overload of communication	Denial of service	Layman	Unnecessary or unlimited access	<== 1 week	Likely	High	Loss of service availability
ENUM query IN SIP or other session server	Limited transport/processing capacity	Overload of communication	Denial of service	Proficient	Moderate	<== 1 week	Possible	High	Loss of service availability
Server keys IN Leaf server	Accessible credentials	Credential manipulation	Manipulation	Expert	Easy	<== 1 month	Possible	High	1. Loss of service availability 2. Data security breaches
DNS records IN Leaf server	Writable cache	Cache poisoning	Manipulation	Expert	Easy	<== 1 week	Possible	High	1. Loss of service availability 2. Data security breaches
Signature on NAPTR IN Leaf server	Writable data records	Credential manipulation	Manipulation	Expert	Easy	<== 1 month	Possible	High	1. Loss of service availability 2. Data security breaches
NAPTR record IN enum core server	Writable data records	Data manipulation	Manipulation	Layman	Moderate	<== 1 month	Possible	High	1. Loss of service availability 2. Data security breaches
Data in transit IN router for enum core server	Limited transport/processing capacity	overload of communication	Denial of service	Layman	Moderate	<== 1 month	Possible	High	Loss of service availability
Data in transit IN router for enum leaf server	Limited transport/processing capacity	Overload of communication	Denial of service	Layman	Moderate	<== 1 month	Possible	High	Loss of service availability
Data in transit IN link from access net to service net	Limited transport/processing capacity	Overload of communication	Denial of service	Layman	Easy	<== 1 month	Likely	Medium	Loss of service availability
Data in transit IN router in access net	Limited transport/processing capacity	Overload of communication	Denial of service	Layman	Easy	<== 1 month	likely	Medium	Local loss of service availability
User credentials in database IN Authentication store (database)	Accessible credentials	Theft of credentials	Interception	Layman	Moderate	<== 1 week	Likely	Medium	Data security breaches

Table 3 24 critical risks found to date

exploited. Essentially, these countermeasures include authentication and authorization of user, integrity protection of signalling messages, and privacy protection of signalling and/or media. More challenging, a means for providing denial of service protection needs to be identified.

The eavesdropping threat applies to signalling data such as authentication information, information about the subscriber ID, or the phone number of the called party. For protection of these various assets the recommended countermeasure is that encryption of user

communication and signalling be applied for outgoing and incoming calls and possibly to the media as well.

To prevent an attacker from registering and using someone else's subscriber-id and authentication information in order to make free calls or even malicious calls, sufficient protection of the authentication information should be provided. For example, in 2G and 3G networks, the subscriber information and authentication keys are safely protected on the USIM card.

For protection against loss of availability of the VoIP service, the situation is more complicated. Some protection can be applied by introducing redundancy and service replication. As unauthorised use of resources can lead to loss of availability for VoIP users, access control mechanisms should be implemented. Additionally, techniques for incident prevention such as Intrusion Detection Systems (IDS) can be used to discover incidents such as denial of service attacks leading to loss of availability of the VoIP service. Indeed, leading IDS vendors are marketing solutions for intrusion prevention designed to protect against attacks targeted at VoIP networks [14]. Strong authentication can mitigate the threat of DoS due to spoofed “bye” packets (as described above).

In order to mitigate the security, availability and privacy risks associated with ENUM, Infrastructure ENUM is being created. In infrastructure ENUM the ENUM data is not published in DNS servers that may be queried from the Internet but the servers and hence the information is only reachable by the SIP servers of the service provider.

The countermeasures addressed in this section will go some way to alleviate some of the security issues in VoIP, however they will not address them all. STF292 is currently extending its TVRA analysis with the impact of countermeasures. Interim results show a similar picture, countermeasures and careful deployment do have an impact but in some cases security may only be brought to sufficient levels by protocol redesign.

Security related issue for VoIP – Handling NAT and Firewalls

One of the challenges facing a VoIP service provider is the widespread use of Network Address Translation (NAT) functionality and firewalls in devices

such as Internet Routers in the subscriber residential network. NATs and firewalls are typically employed to address generic networking security issues of access control and topology hiding.

However these NATs and firewalls cause issues while deploying VoIP as they make it difficult for incoming calls to be received by a terminal behind the NAT or firewall. Both can also have a detrimental effect on QoS. The main problem is that NAT only handles outgoing connections, which means essentially, that incoming calls will be dropped by the NAT unless a solution is applied. Furthermore, the end-to-end SIP messaging between the clients (SIP User Agents) contains details of the private IP addresses and ports that the User Agents intend to use for media flows. The problem when there is a NAT between two such endpoints is that the User Agent attempts to use these private addresses and port numbers to send/receive media, and the connection fails, as these cannot be routed in the public address space.

VoIP protocols are designed in such a way that the signalling is handled by one protocol, and another handles the media. To make matters worse, the port on which the media is sent is chosen randomly. For these protocols to pass through a firewall, the specific static and the range of dynamic ports must be opened for all traffic. If the firewall is set to open up for any port in this range, this means a huge vulnerability which would not be good firewall policy.

There are several different proposals for working around the NAT (and Firewall) problem. In this paper we present five solutions. The task at hand for this work around solutions is to discover the public IP address and port that is used by the NAT, and keep the binding valid so that it can be used for incoming and outgoing calls.

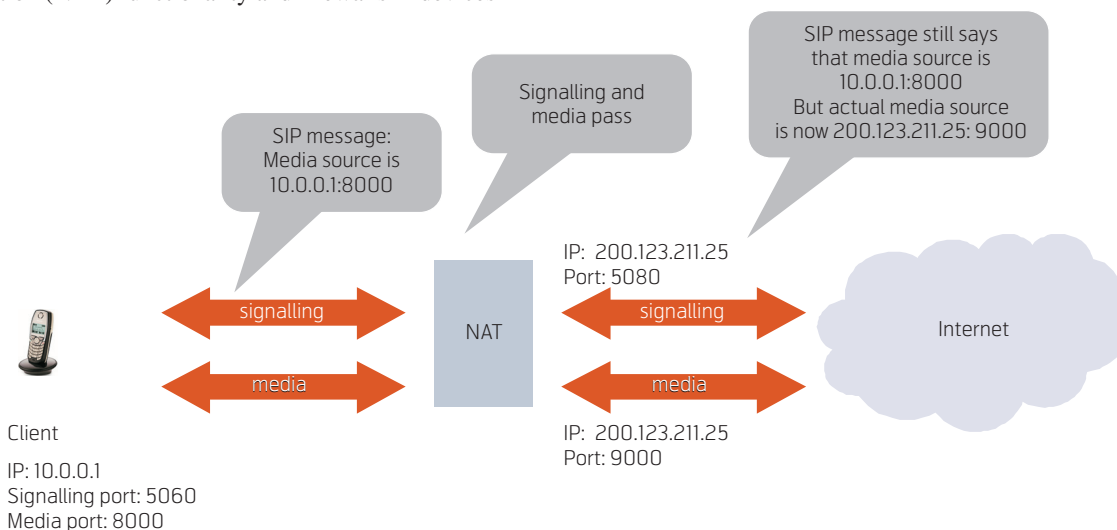


Figure 6 NAT breaks the end-to-end media flow

Most of these proposals achieve this by creating a binding (RTP over UDP) to the SIP server that is held open over UDP.

While this solution is satisfactory, there are some threats associated with it; in some cases, it may be possible to fake packets which cause the terminal to ring out even though the subscriber is not there, or worse, can send a “bye” causing the call to be disrupted.

STUN

STUN is the acronym for “Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)”, the NAT traversal solution specified in the IETF RFC3489 [15]. STUN allows entities behind a NAT to first discover the presence of a NAT and the type of NAT, and then to learn the address bindings allocated by the NAT. STUN is a simple client-server protocol. The STUN client asks the STUN server to determine the external IP address and port of NAT, and the NAT binding is subsequently refreshed (kept open) by STUN messages. The STUN client then sends a call request to the SIP proxy containing the public address and port of the NAT. Some drawbacks of this solution are that it can only be used with some NATs that open predictable ports on subsequent mappings. This renders the information provided by the STUN server useless for initiating communication to other addresses than the STUN server address. The NATs that do support STUN are vulnerable to port scans, and it does not support TCP based SIP devices. As such STUN utilizes a security weakness in existing NATs, which hopefully will be solved soon by the vendors of these devices.

IGD

The increasingly popular UPnP [16] Internet Gateway Device (IGD) protocol, when implemented in a NAT router, allows a terminal or PC behind that router to explicitly request IP ports to be routed to it. This allows a SIP User Agent to send the correct signalling as if it had the router’s public IP address to itself. This approach in itself is functional but also a big security risk. When a PC may request ports to itself to be opened this allows a virus entering the system by other means (e.g. email) to open a port to itself so the computer becomes controllable from the Internet by a remote hacker. This proposed approach is therefore also of questionable merit.

ALG, B2BUA and SBC

This section deals with three very similar concepts, which unfortunately have radically different names. What they all have in common is that an application-aware element is joined with a packet or media forwarding engine so that this engine can receive direc-

tions from the application element on which flows to let pass and which not to, while the application element may learn from the packet engine which IP address ports to use on the outgoing flows.

The first concept addressed in this fashion is the Application Level Gateway (ALG). The ALG is a function embedded in the NAT or firewall *typically at the customer premises* that understands (in this case) the VoIP protocol used. This ALG may use its closeness to the NAT function to synchronize the port mapping and to place the correct media ports and addresses in the outward signalling. When the same function is placed in the network of a service provider it is called a Back to Back User Agent (B2BUA) or Session Border Controller (SBC). With the B2BUA, two user agents are placed back to back in a device between the client (user agent) and the SIP proxy, taking what is traditionally a SIP end-to-end call and mediating it through a central SIP server. The B2BUA communicates with the packet forwarding engine that relays the media traffic. The primary function is to replace private addresses with public routable addresses so that the media can be routed through the networks (both public and private) to reach the client devices. The B2BUA enables service providers to maintain call state from beginning to completion of the call.

The Session Border Controller (SBC) performs the same functions as above and is also typically an entity enforcing the access policy of the service provider, hence the word “controller” in the name. Session Border Controllers typically perform a thorough inspection of the messages they pass through and have knowledge of the services they allow to pass through. They may also explicitly check with policy entities in their network to obtain authorization for user-initiated sessions. Several vendors of SBCs are expanding their products to support control over all kinds of protocols and communication, including that of peer to peer services like Skype. Although SBCs provide a functional solution they seem to duplicate work that also will be done by SIP proxies in the network. They suffer from the issues inherent in ALG-based solutions in that they need to be compatible with the protocols they support (which in the case of SIP leads to a whole range of variants that need to be supported) and they may have scalability problems.

Middlebox

A problem with the ALG approach is that it requires the version of the VoIP protocol used (e.g. SIP) to be compatible with the one understood by the ALG. As there are many different flavours of SIP in use this may prove to be a problem in practice. The Middle-

box [17] approach splits the ALG and the NAT function. It allows many application servers to be deployed in the VoIP service provider's network for the many signalling protocols and services in use while all of them have a single interface to the NAT functions at the edges of the network [18] [19] [20].

NAT and IPsec

IP security (IPsec) is used e.g. by 3GPP IMS to encrypt the SIP signalling and provide a secure and transparent tunnel through which the media may flow unimpeded. However, NAT and IPsec were not designed to work together, and therefore, a solution has been created by the IETF so that IPsec can traverse NATs. If IPsec is used for authentication, the modification of the IP packet by NAT causes a failed integrity check for IPsec. Therefore, NAT traversal cannot be defined for IPsec with the IPsec authentication header (AH). The IETF has published NAT Traversal (NAT-T) [21] [22] [23], which aims to address these issues using a method for encapsulating IPsec ESP packets into UDP packets for passing through routers or firewalls employing Network Address Translation (NAT). However, sending IPsec through NATs may only be applicable for signalling. Because IPsec packets are much larger than regular VoIP media packets, securing VoIP media with IPsec will increase the transmission delay and bandwidth needed for the media. Although this will not be a problem in an enterprise network, increased delay and bandwidth needs are likely to significantly reduce the media quality on real-life broadband access networks such as ADSL.

Media reflection

The TURN protocol was once proposed to the IETF to allow communication of a SIP client to a packet reflector service on the Internet. This would implement a media relay for SIP end-points behind a NAT. The approach however is not ideal. It assumes the clients have a trust relationship with a TURN server and request session allocation based on shared credentials. This has scalability issues, requires complex changes in the SIP clients, as the TURN protocol is difficult to implement, has no possibility of distributing the load and complicates the configuration of the SIP user agent. The TURN protocol seems to be no longer developed.

Another approach, which requires no changes in the SIP devices, is to reuse the trust relationship the SIP device already has with the SIP Proxy. In contrast with how TURN works, the SIP Proxy and not the User agent does the session reservation for the media relay. This has the immediate advantage that the SIP UA does not have to have any TURN capability built in, and secondary a database with user credentials

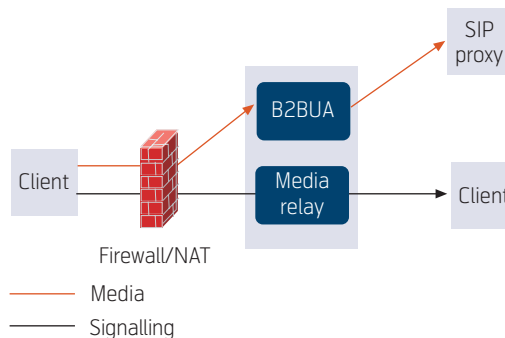


Figure 7 Back to back agent assisted NAT traversal

does not need to be stored on both the TURN server and the client. Another advantage is the fact that the SIP Proxy always has more clues about where the best place is to assign a media relay for a SIP session apart from the SIP devices themselves. This allows per call allocation of a media relay session in an optimum place on the Internet and solves the load balancing and scalability of the media relay function. In this approach the SIP proxy indicates to the client the reflector IP address and port, and subsequently the SIP terminal is required to initiate the media flow and send this to the media reflector in the network. This approach is employed by a number of SIP clients embedded in ATAs and handsets. This does however require a SIP proxy server and bandwidth to be available to perform the reflection.

NAT Conclusion

The NAT and Firewall issues are a major problem for the widespread deployment of VoIP. Unfortunately, there is no complete solution to problems imposed by the use of NATs and Firewalls. This is due in part to the wide variety of types and implementations avail-

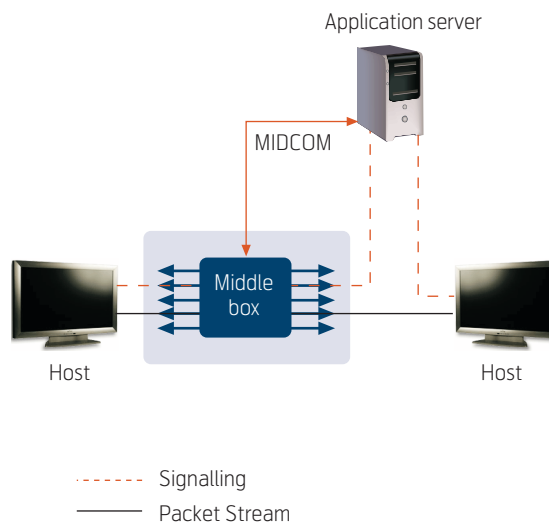


Figure 8 Middlebox control

able and configured in the market today. If the operator owns the equipment, however, then the operator can preconfigure the NAT and/or Firewall so that it will work with the operator's SIP servers. An operator may choose to provide the router to the customer preconfigured to ensure that the NAT works with the VoIP implementation.

Standardization

ITU-T

Standardisation of VoIP protocols in the ITU-T is the responsibility of Study Group 16. The security architecture for the H.323 protocol suite is provided in H.235 version 2. Currently, SG16 is working on the development of a framework and roadmaps for the harmonized and coordinated development of multimedia telecommunication standardization over wired and wireless networks, and in particular, security of multimedia systems and services [24].

IETF

The Session Initiation Protocol (SIP) working group carries out further development of SIP in the IETF [25]. The following is a list of some of the IETF RFCs specifying security for SIP:

- Network layer security: IPsec (RFC 2401) [26];
- Transport layer security: TLS (RFC 2246) [27].
Note: This provides transport layer security over connection-oriented protocols; for SIP, this means TLS over TCP. TLS cannot run over UDP as TLS requires a connection-oriented underlying transport protocol;
- S/Mime (RFC 2633) [28] for e.g. protection of SIP signaling;
- Authenticated Identity Body (AIB) Format (RFC 3893) [29] for e.g. authentication of the ID of the sender;
- Privacy mechanisms (RFC 3233) [30] to protect personal identity information.

A Complete list is available at the official SIP charter [25].

It should be noted that although the status quo for SIP implementations on the market today is to use the MD5 algorithm for digest authentication, the SIP standard, RFC 3261, does not mandate the use of MD5 [31]. Any algorithm may be implemented, as in the case for H.235, making it possible to support stronger authentication in future implementations.

VoIP Security Alliance – VoIPSA

VoIPSA, the recently formed Voice over IP Security Alliance, is a collaborative initiative by VoIP Information Security vendors, and providers. The aim of the alliance is to address VoIP security related issues in order to mitigate VoIP security risks by promoting VoIP security research, spread educational information on VoIP security and awareness as well as making VoIP testing methodologies and tools freely available [18].

ETSI TISPAN

TISPAN (= Telecommunication and Internet converged Services and Protocols for Advanced Networking) is the ETSI technical body on next generation networks (NGN) and addresses convergence of fixed and wireless networks. The NGN will provide a multi-service, multi-protocol, multi-access, IP based network. ETSI TISPAN NGN Release 1 uses "core" IMS as one of the NGN architecture components (as there is an agreement on reuse of 3GPP/ 3GPP2 IMS in comprehensive NGN plans). TISPAN is defining an IP-based PSTN/ISDN replacement service, the PSTN/ISDN Emulation Service (PES), which is identical to the PSTN/ISDN Telephony service but offered over an IP infrastructure in order to enable the use of the Existing ISDN Supplementary services. TISPAN is also defining a Voice Service (Simulation), which is similar but not identical to the existing PSTN service including "important" "supplementary" services but adapted to the 3GPP IMS environment. SIP extensions that are needed by TISPAN are expected to be included in 3GPP R7 IMS capabilities. Figure 9 shows the overall TISPAN NGN functional architecture.

The NGN is an enabler for Service Providers to offer real-time and non real-time communication services between peers or in a client-server configuration allowing nomadicity and mobility of both users and devices. There is a strong emphasis on security on a managed IP network and on regulatory compliance on issues such as Lawful Intercept, Number portability, and Emergency services.

TISPAN Working Group (WG) 7 is responsible for the management and co-ordination of the development of security specifications for the NGN project. For TISPAN NGN Release 1 and beyond, TISPAN WG7, assisted by the specialists task force (STF) 292, is defining the security requirements for the NGN, the security architecture for the NGN, conducting threat and risk analyses for the NGN scenarios and in particular, TISPAN WG7 and the STF 292 are proposing countermeasures related to the security issues presented in this paper.

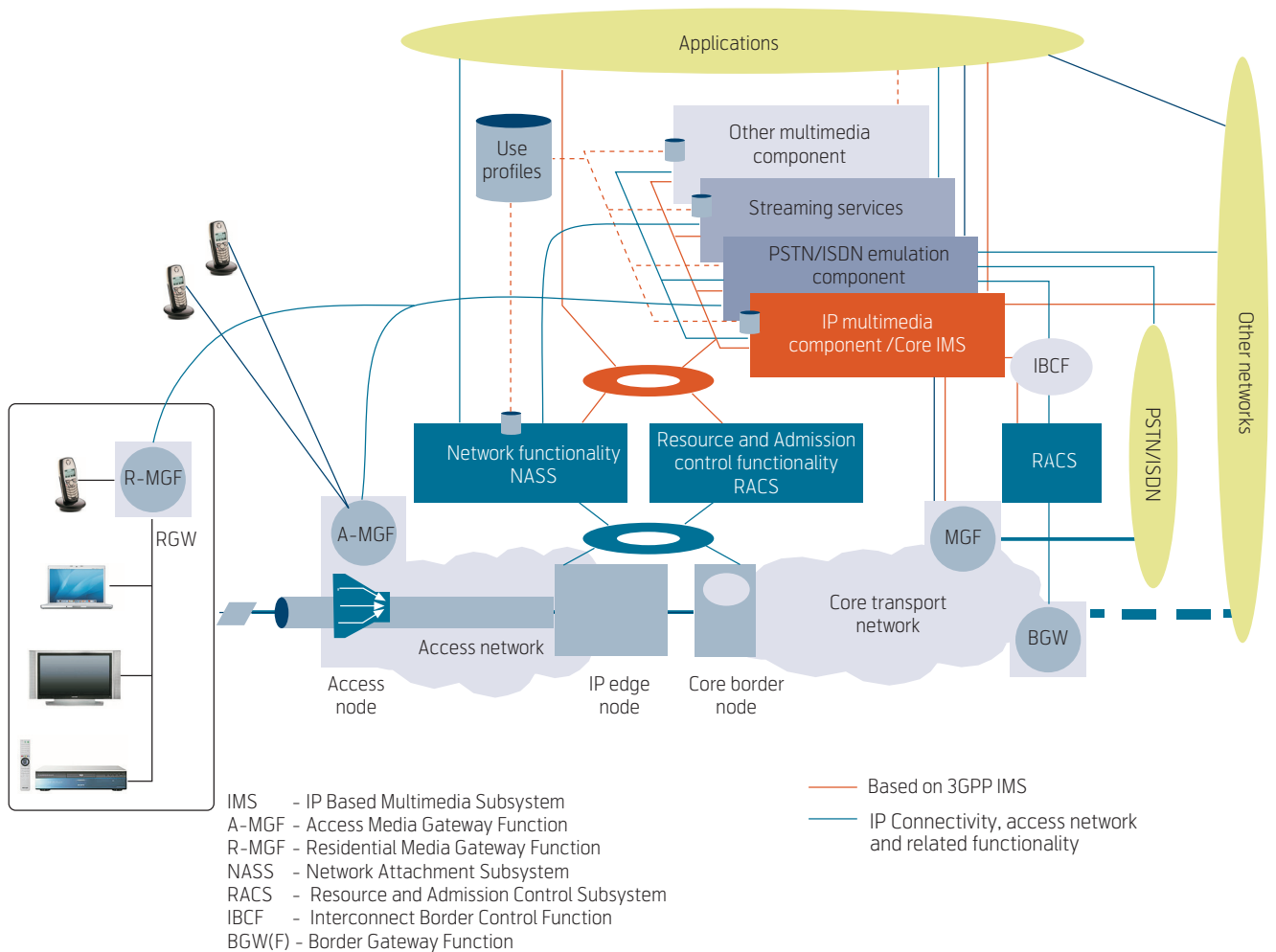


Figure 9 TISPAN NGN overall architecture

Non standard solutions

For an example of a non-standard peer-to-peer VoIP application, we consider Skype, a peer-to-peer VoIP application available worldwide [32]. Like other VoIP applications available today, Skype authentication is username and password based, and is therefore susceptible to the same types of attacks and social engineering techniques for stealing usernames and passwords. And, as for any other VoIP software client it is vulnerable to manipulation via malicious software on the terminal, and the first incidents of potentially serious security vulnerabilities have already been reported [33]. However, unlike PSTN/ISDN and other VoIP applications available today, Skype media and voice communications are encrypted so that Skype communications are protected against eavesdropping. Skype protocols are proprietary and secret, making it difficult to conduct a security analysis of Skype, while reverse engineering can produce some indications about how Skype works as in [34].

Conclusions/Summary

This paper has discussed issues and the threats to VoIP and countermeasures to mitigate these threats. Early results from the threat analysis as is being done in TISPAN WG7 STF292 [35], have shown that there are many security risks in deploying VoIP technology today. There is an increasing awareness of the potential problems and there are initiatives working to improve VoIP security. Indeed, there are technically feasible solutions, yet market constraints determine which solutions are implemented. It is therefore crucial to conduct risk analysis both at a service provider and network level, but also at a broader level, as is being done in TISPAN WG7 STF 292 [20].

The true test of whether VoIP implementations are robust enough and provide sufficient security measures will come when VoIP essentially replaces the PSTN/ISDN and GSM voice services. In the meantime, it is extremely important to rigorously test security of VoIP implementations to both test against resilience to known vulnerabilities as well as pinpointing unknown vulnerabilities and resolving other

security issues. Should these risks be underestimated we could see the reliability of the telephone service fall dramatically.

Acknowledgements

The research on which this paper reports has been funded in part by the Research Council of Norway project SARDAS (152952/431).

Thanks to colleagues at Telenor, Dole Tandberg, Bernt Haram, and Tor Hjalmar Johannessen for interesting discussions on security issues for VoIP. Thanks also to Tor Hjalmar Johannessen and Martin Sierink for commenting on earlier versions of this article.

References

- 1 *Stadig flere tar i bruk bredbånd og bredbåndstelefoner*. April 9, 2006 [online] – URL: http://www.npt.no/portal/page?_pageid=121,47056&_dad=portal&_schema=PORTAL&p_d_i=-121&p_d_c=&p_d_v=45729
- 2 *OnInstant*. News release, IT and Telecoms industry legend joins OnInstant. May 2005 [online] – URL: www.on-instant.com.
- 3 *The History of Hacking*. May 2005 [online] – URL: <http://www.roadnews.com/html/Articles/historyofhacking.htm>.
- 4 FCC News Media Information 202 / 418-0500. *Commission Requires Interconnected VoIP Providers to Provide Enhanced 911 Service*. May 2005 [online] – URL: <http://www.fcc.gov/>.
- 5 Norwegian Post and Telecommunication Authority. *Regulering av bredbåndstelefoner*. Published April 15, 2005. URL: <http://www.npt.no>.
- 6 Jensen, W. VoIP – Regulatory aspects from a Norwegian perspective. *Teletronikk*, 102 (1), 23–26, 2006. (This issue)
- 7 Franks, J, Hallam-Baker, P, Hostetler, J, Lawrence, S, Leach, P, Luotonen, A, Stewart, L. *HTTP Authentication: Basic and Digest Access Authentication*. June 1999. (RFC 2617)
- 8 Rivest, R. *The MD5 Message-Digest Algorithm*. April 1992. (RFC 1321)
- 9 CERT. *CERT Advisory CA-1998-01 Smurf IP Denial-of-Service Attacks*. March 2004 [online] – URL: <http://www.cert.org/advisories/CA-1998-01.html>.
- 10 Leyden, J. *Say hello to the Skype Trojan*. Published October 18, 2005. URL: http://www.theregister.co.uk/2005/10/18/skype_trojan/.
- 11 Ulseth, T, Stafnes, F. Real-time communication on IP networks. *Teletronikk*, 102 (1), 3–22, 2006. (This issue)
- 12 Ulseth, T. Telephone Number Mapping (ENUM) – A short overview. *Teletronikk*, 102 (1), 40–42, 2006. (This issue)
- 13 ETSI. *Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); 3G security; Access security for IP-based services*, 2005. (ETSI TS 133 203)
- 14 *NFR Security Launches VoIP Protection Package for Sentivist Intrusion Prevention Solution*. March 9, 2006 [online] – URL: <http://www.nfr.net/news/detail.php?id=305>
- 15 Rosenberg, J, Weinberger, J, Huitema, Mahy, C R. *STUN – Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)*. March 2003. (RFC 3489)
- 16 *UPnP Forum*. April 10, 2006 [online] – URL: <http://www.upnp.org/>
- 17 van Willigenburg, W, de Boer, M, van der Gaast, S. Middleboxes: Controllable media firewalls. *Bell Labs Technical Journal*, spring 2002.
- 18 *VoIP Security Alliance*. March 9, 2006 [online] – URL: <http://www.voipsa.org/>
- 19 *PROTOS Test-Suite: c07-sip*. March 9, 2006 [online] – URL: <http://www.ee.oulu.fi/research/ouspg/protos/testing/c07/sip/>
- 20 *Specialist Task Force 292: TISPAN security: Standards development in support of the eEurope secure and trusted network environment*. March 9, 2006 [online] – URL: <http://portal.etsi.org/STFs/TISPAN/STF292.asp>
- 21 Kivinen, T, Swander, B, Huttunen, A, Volpe, V. *Negotiation of NAT-Traversal in the IKE*. January 2005. (RFC 3947)
- 22 Aboba, B, Dixon, W. *IPsec-Network Address Translation (NAT) Compatibility Requirements*. March 2004. (RFC 3715)

- 23 Huttunen, A, Swander, B, Volpe, V, DiBurro, L, Stenberg, M. *UDP Encapsulation of IPsec ESP Packets*. January 2005. (RFC 3948)
- 24 ITU-T. *ITU-T Study Group 16*. May 2005 [online] – URL: <http://www.itu.int/ITU-T/study-groups/com16/area.html>.
- 25 *The Session Initiation Protocol (SIP) working group*. June 2005 [online] – URL: <http://www.ietf.org/html.charters/sip-charter.html>.
- 26 Kent, S, Atkinson, R. *Security Architecture for the Internet Protocol*. November 1998. (RFC 2401)
- 27 Dierks, T, Allen, C. *The TLS Protocol Version 1.0*. January 1999. (RFC 2246)
- 28 Ramsdell, B. *S/MIME Version 3 Message Specification*. June 1999. (RFC 2633)
- 29 Peterson, J. *Session Initiation Protocol (SIP) Authenticated Identity Body (AIB) Format*. September 2004. (RFC 3893)
- 30 Petersen, J. *Privacy Mechanism for SIP*. November 2002. (RFC 3323)
- 31 Rosenberg, J, Schulzrinne, H, Camarillo, G, Johnston, A, Peterson, J, Sparks, R, Handley, M, Schooler, E. *SIP: Session Initiation Protocol*. June 2002. (RFC 3261)
- 32 *Skype*. March 9, 2006 [online] – URL: <http://www.skype.com/>
- 33 Leyden, J. *Scramble to fix Skype security bug*. Published October 25, 2005. URL: http://www.theregister.co.uk/2005/10/25/skype_vuln/
- 34 Biondi, P, Desclaux, F. *Silver Needle in the Skype*. April 18, 2006 [online] – URL: <http://www.blackhat.com/presentations/bh-europe-06/bh-eu-06-biondi/bh-eu-06-biondi-up.pdf>
- 35 ETSI, Specialist Task Force 292. *TISPAN security: Standards development in support of the eEurope secure and trusted network environment*. April 18, 2006 [online] – URL: <http://portal.etsi.org/stfs/tispan/STF292.asp>
- 36 ETSI. *TIPHON Release 4; H.248/MEGACO Profile for TIPHON reference point I3 Part 2 – ICF Control over Reference Point I3*. April 2002. (ETSI TS 102 108-2)
- 37 ETSI. *Telecommunications and Internet Protocol Harmonization over Networks (TIPHON) Release 4; Interface Protocol Definition; TIPHON Extended H.248/MEGACO Package (EMP) Specification Part 2 – ICF Control over Reference Point I3*. April 2002. (ETSI TS 101 332-2)

Judith E.Y. Rossebø is a Research Scientist at Telenor R&D. Prior to joining Telenor in October 2000 she worked three years as a systems engineer at Alcatel Telecom Norway and one year as an assistant professor teaching mathematics at the University of Tromsø. At Alcatel she worked with IN, and dimensioning, performance, dependability and traffic control in telecommunication networks. She received a cand.scient. degree in Mathematics from the University of Oslo in 1994 and is currently working on a PhD at Norwegian University of Science and Technology (NTNU), department of Telematics, in the SARDAS project. Since January 2003 she has been the Chairman of ETSI TISPAN WG7 Security. Her research interests include security in general; security issues in multimedia communications services, and in particular securing availability of services.

email: Judith.Rossebo@telenor.com

Paul Sijben is chief technologist of EemValley Technology. In 1993 he graduated from the University of Twente in computer science and stayed there to do research on multimedia operating systems. In 1997 he joined Lucent Bell-Labs in the Netherlands and left in 2002 as a Distinguished Member of Technical Staff working on Voice over IP research and standards (accomplishments include substantial work on specification of the TIPHON architecture and the H.248/MEGACO protocol) [35][36]. In 2002 Paul became CTO of PicoPoint, a Wi-Fi hotspot back office and roaming pioneer. In 2004 he left PicoPoint to found EemValley Technology to create the next generation of open and secure telecommunication infrastructure and services. Paul is currently also active in the ETSI TISPAN Specialist Task Forces on next generation services and security.

email: sijben@eemvalley.com

The Telenor Nordic Research Prize 2005

TERJE ORMHAUG



Terje Ormhaug is Senior Research Scientist at Telenor R&D and Secretary for the Research Prize

Among candidates from Denmark, Finland, Norway and Sweden who were nominated for the Telenor Nordic Research Prize 2005, the Jury unanimously chose professor Ramjee Prasad at the University of Aalborg, Denmark, as the winner. The theme for the 2005 prize was:

“Enabling technologies and communication based value added services for the home, leisure and professional environments”

The significance of the theme for Telenor is that personalised services and universal access adapted to available terminals, will be a vision for the future development of our industry. The research cited could include aspects of technology, services, user acceptability and business opportunities.

Professor Prasad is a world-class researcher with a long history of impressive merits. Among his noteworthy achievements are that he has been author or co-author of 16 books and has a list of more than 500 publications. He has been an active creator of academic and industrial networks internationally, and has been very active in dissemination of research findings. Besides his own work he has made a great effort to promote research in Aalborg, for example by organizing international programs involving both universities and industries.

In particular for the theme of this year, professor Prasad has been an early leader to formulate a vision around Personal Area Networks. This is judged to be a key concept for flexible access to multimedia and personalized services. His work has in general had significant impact on technology development, constituency building, and concept formation, and future impact on applications is expected. He is a leader in an important but still formative field – 4G mobile communications.

The Telenor Nordic Research Prize has been granted since 1997, with the motivation to underline the importance of research and innovation in information technology and telecommunications, and to support Nordic efforts in this area. The prize consists of a diploma and NOK 250,000, and may be given for work within a broad range of professional fields. These may encompass areas like technology, networks and service development, economy and markets, and user behaviour, as well as new areas that may appear to be of importance in the future.

For 2006 the theme will be about systems and organisational aspects of emergency communication. According to the statutes the prize can be awarded to individuals or research groups for work that has been carried out in the Nordic countries, or by citizens from these countries. Everyone may nominate candidates for the prize. The deadline for proposing candidates will be early August. The exact theme description and deadline, along with criteria and procedures for entering proposals, will be made public on: <http://www.telenor.com/rd/ra/index.shtml>

Themes and winners of previous years have been:

- 1997 *Video Coding*
Dr. Gisle Bjøntegård (N)
SW tools and languages
Birger Møller-Pedersen, Dag Belsnes and Øistein Haugen (N/DK)
- 1998 *Development of internet technology*
Professor Stephen Pink (S)
- 1999 *Advanced applications of communications and information services*
The Telepathology group (N)
- 2000 *Enabling technologies for advanced ICT systems and services*
Professor Peter Andreksson (S)
- 2001 *Research in socio-economic impact of ICT*
Professor Jon Bing (N)
- 2002 *Mobility and wireless access to Internet. Technologies, new services and applications*
Professor Christian S. Jensen (DK)
- 2003 *Technologies and systems enabling new communication services*
Dr. Haakon Bryhni (N)
- 2004 *SW of importance for the creation or improvement of communication services*
Professor Claes Wohlin (S)

The members of the jury are:

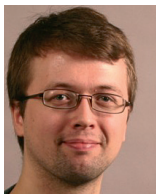
- Vice President *Berit Svendsen* (leader), Telenor Nordic Fixed, Norway
- Professor *Peter T. Kirstein*, University College London
- *Ekonomie Dr. Erik Bohlin*, Chalmers University of Technology, Sweden
- Associate Professor *Mads Christoffersen*, The Technical University of Denmark
- Professor *Gunnar Stette*, The Norwegian University of Science and Technology
- Professor *Olli Martikainen*, The University of Oulu, Finland

The Unpredictable Future: Personal Networks Paving Towards 4G

RAMJEE PRASAD AND RASMUS L. OLSEN



Ramjee Prasad is Director of Center for Teleinfrastruktur (CTIF) at Aalborg University, Denmark



Rasmus L. Olsen is a PhD student at Aalborg University, Denmark

In this paper we discuss how the network paradigm Personal Networks will become an evolutionary and revolutionary step for communication technology towards the fourth generation communication (4G). For 4G, not only higher data rates, user capacity, latency and data coverage are parameters of interest, but also technology convergence, personalisation and security play leading roles. In this paper we describe and discuss how Personal Network addresses exactly these issues, for which it will pave the technology development towards the envisioned 4G.

1 Introduction

Within the last two or three decades the means of communication has undergone a dramatic development, and created a number of technologies which enable communication between devices. Figure 1 illustrates how the progress towards the next generation in communication technology, 4G, can be perceived as a tree with many branches.

The diversity of technologies is the result of many factors, e.g. technical problems in different domains, ranging from the physical to the application layer, various market players and their interests and so on. A set of examples in the differences that can be found in the wireless standards today are:

- Coverage
- Data-rates
- Services
- Medium Access Control protocols
- Quality of Service methods

- Network architecture
- Mobility solutions
- Security methods: Authentication, key-management, encryption schemes etc.

The paper is divided into six sections: The first section provides the background for current technology and why convergence and personalisation are such important aspects in 4G. In the second section, we introduce Personal Networks where we define and describe the overall picture of this new network paradigm. In the third section, we describe the construction and some of the most important technical challenges for Personal Networks that are to be overcome to achieve the development of this technology. In section four we introduce many of the security and privacy issues that must be overcome before a Personal Network can be realised. In section five we provide three cases of how Personal Networks specifically address personalisation and how it can adapt to the user and the user's current needs. Finally, in section six, we conclude and provide an outlook for this concept.

1.1 Evolution of mobile technology

Traditional requirements to a next generation communication systems have typically involved *increased user capacity, low latency, higher data rates and coverage*, e.g. [2]. Figure 2 shows how the technology has progressed from 2G cellular systems towards existing ones, and finally how this in a natural way leads towards the requirements for 4G.

However, with the introduction of transportation of not only voice data, but also other types of data over wireless networks, *users, applications and the business market* have become important players. This pushes the traditional view of 4G as just another generation of networks offering even higher data rates etc., to the limit. For the business market, and ultimately the end user, it is important that the communication network has a *low deployment cost*. For the user, it is important that the technology is *person-*

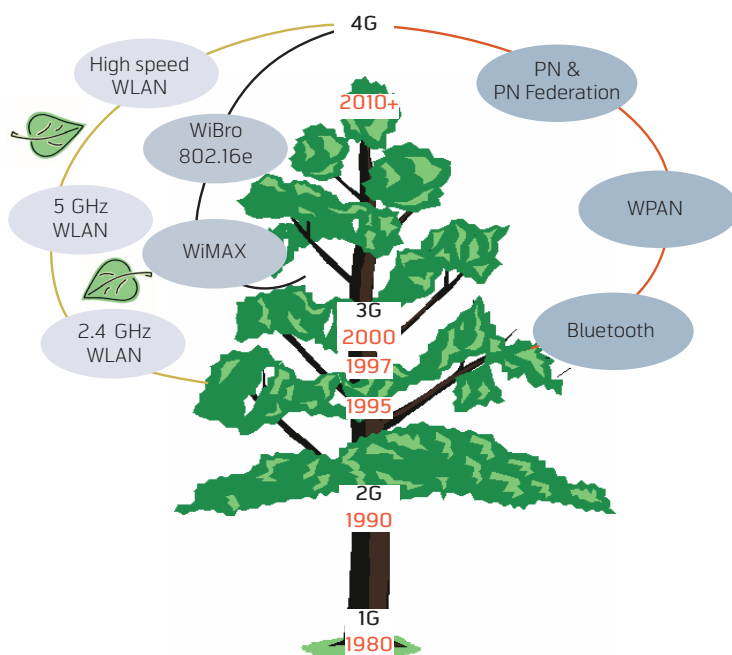


Figure 1 The progress tree for communication technology [1]

alised and secure, since the user will eventually use more personalised services and applications, and therefore these two parameters will become the main focus of the end user. Furthermore, with the mobility that exists in the world today, e.g. people travelling around the world everyday, we still need to interact with services and applications at home, in our office, around us, etc. For this reason, global connectivity and an adaptive network structure which is able to overcome the heterogeneity that exists today, will become necessary.

1.2 Convergence toward 4G

In this paper we will define 4G as an evolutionary and revolutionary new fully IP-based integrated system of systems and network of networks achieved after convergence of wired and wireless networks as well as computers, consumer electronics, and communication technology and several other convergences that will be capable of providing 100 Mb/s and 1 Gb/s, respectively in outdoor and indoor environments, with demand-driven end-to-end QoS and high security, offering any kind of services at any time as per user requirements, anywhere with seamless interoperability, always on, at an affordable cost, with one billing and fully personalized.

As stated, convergence of technology is really a needed part for achieving 4G. However, convergence can happen on many levels, and we briefly describe two examples in the following sub sections.

1.2.1 Terminal convergence

At the terminal level, which is the user's first and closest experience with the technology, convergence can be seen as an invisible and seamless service provisioning, in the sense that apparently to the user, the devices can easily interact with each other and offer services to the user as he/she needs under given circumstances. An example is the *flying screen* [1], illustrated in Figure 3.

This concept captures the user's need for a display service, which can be either the TV, laptop screen or the mobile phone. The vision is that any content can be shown at any time and anywhere. However, for today's technology it is not a simple matter to show pictures taken by the camera on the mobile on the TV if the user wishes to show holiday pictures to his/her friend. With the flying screen, the holiday pictures are easily shown on either of the displays, without the user having to perform all kinds of technical setups and software installations for interacting with the service providing device.

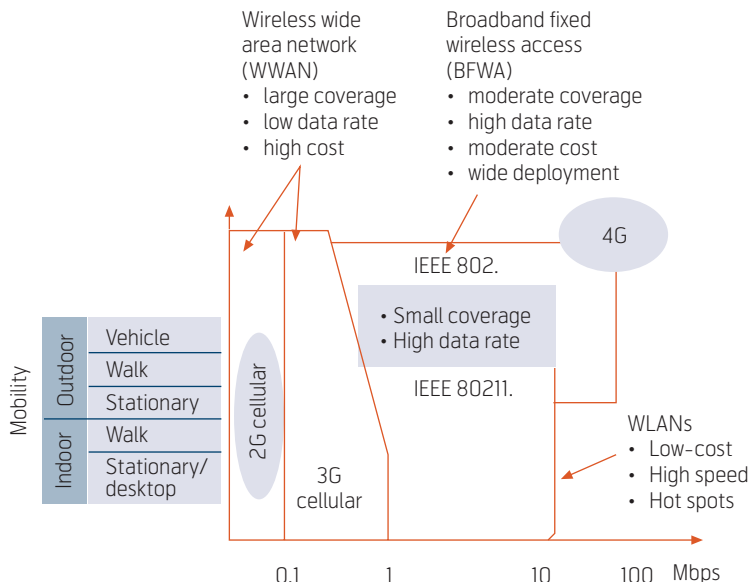


Figure 2 Progression of wireless technology [3]

1.2.2 Network convergence

Today a wide range of network technologies already exist on the market, some are IP based, and others are not, e.g. Bluetooth, Zigbee. Some network are local such as Wireless Local Area Networks (WLAN), Wireless Personal Area Networks (WPAN), while others are global, e.g. the Internet, satellite networks. From a user's point of view, a common technology is beneficial, since this basically enables communication between various devices, over short as well as long distances through whatever communication means that is available. The trend at network level is

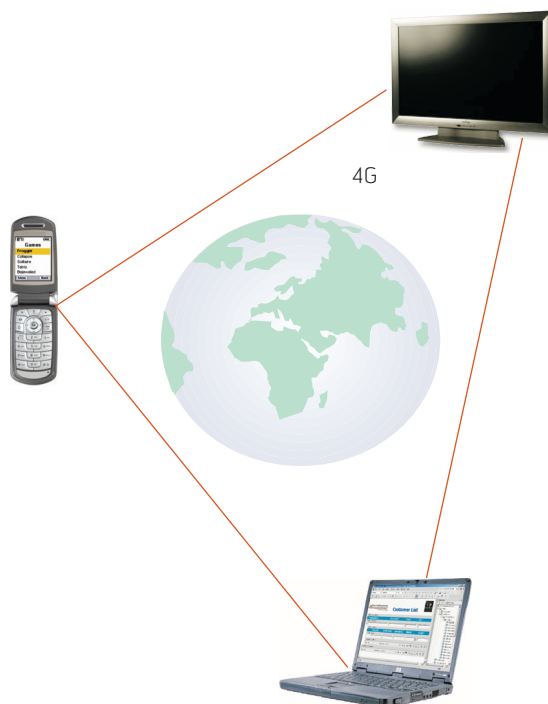


Figure 3 The flying screen concept [1]

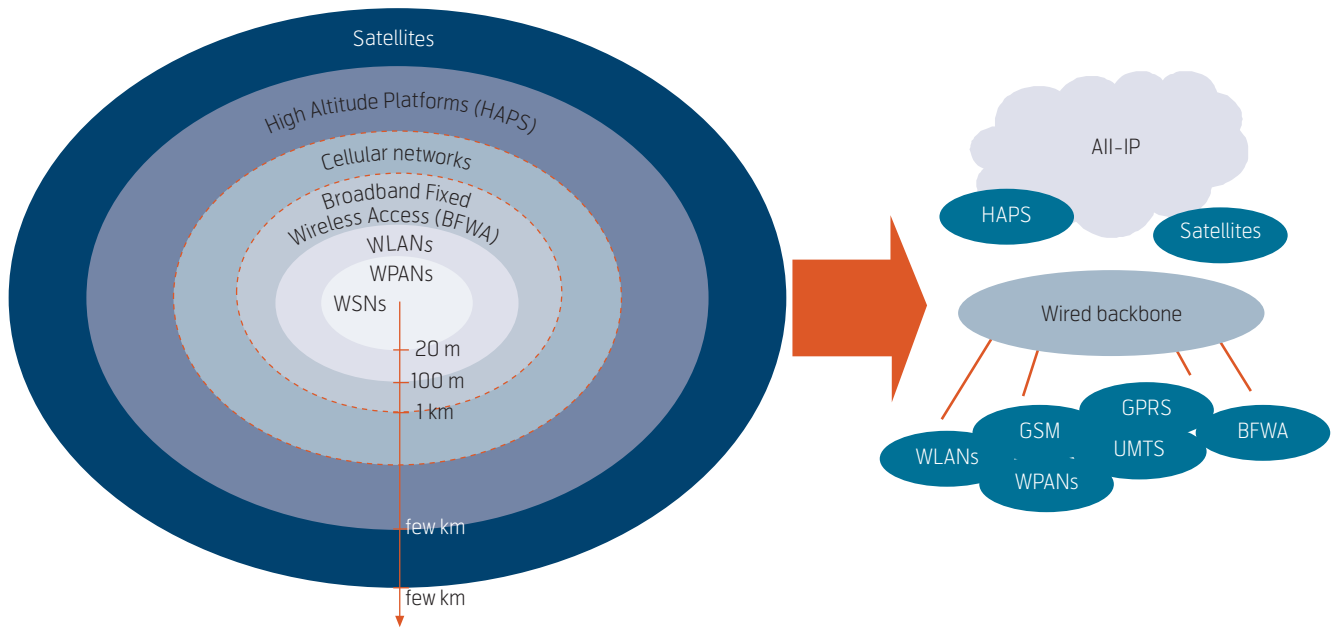


Figure 4 Convergence in network technology [4]

that an IP based solution will become the major technology for future networks [4] and is illustrated in Figure 4.

Furthermore, a common platform such as IP makes software development much easier, for new network and application components, but also for services. This is a key issue that eventually will benefit the user, as a standardised interface will make interaction between applications much easier, and ultimately will make the convergence on terminal level happen.

1.3 Personalisation and personalised services

Even in today's market, the end user plays an important role, but for 4G the user will be the centre point and not the technology as it has been previously. In this sense, user centricity means applications and services will be developed with the end user as a person and not some anonymous entity that will have to use whatever the technology is capable of offering.

This person centric approach means that applications and services will need to adapt to who the user is, the user's interests and current situations. From a user's point of view, some high level requirements can be set, to which the technology simply has to adapt:

- Being able to use a service or application anywhere at any time;
- Doing this in a cheap and efficient way;
- Will not need to do a lot of technical parameter setup herself;
- The user's current situation and interests.

In fact, meeting all these requirements is not only a matter of ensuring high coverage, high data rates etc., but is to a large degree also a matter of taking into account *user profiles*, *user context* and *adaptation* towards services and applications that the user will be using. Taking into account these matters enables the technology to adapt its behaviour by changing system parameters, making intelligent selections on e.g. air interfaces or filtering information not relevant to the user, e.g. context aware service discovery [5].

Figure 5 gives an overview of what personalisation constitutes, and how this concept will undergo a transformation from simple and single research topics and move towards a convergence for where the technology must closely cooperate in order to achieve the envisioned intelligent behaviour of 4G systems.

The envisioned personalisation will have a potential impact on our society and the way we communicate, as it will assist our everyday of interacting with our work, family, friends and so on. It is also this concept that acts as a trigger to develop a new network paradigm such as Personal Network.

All in all, it is easily seen that the progress towards 4G is not only about achieving better data rates, lower latency and increased user capacity as it would have been if 4G followed the same development progression as the previous technologies; but it is also about the incorporation of the user to a much higher degree than previous generations of communication networks.

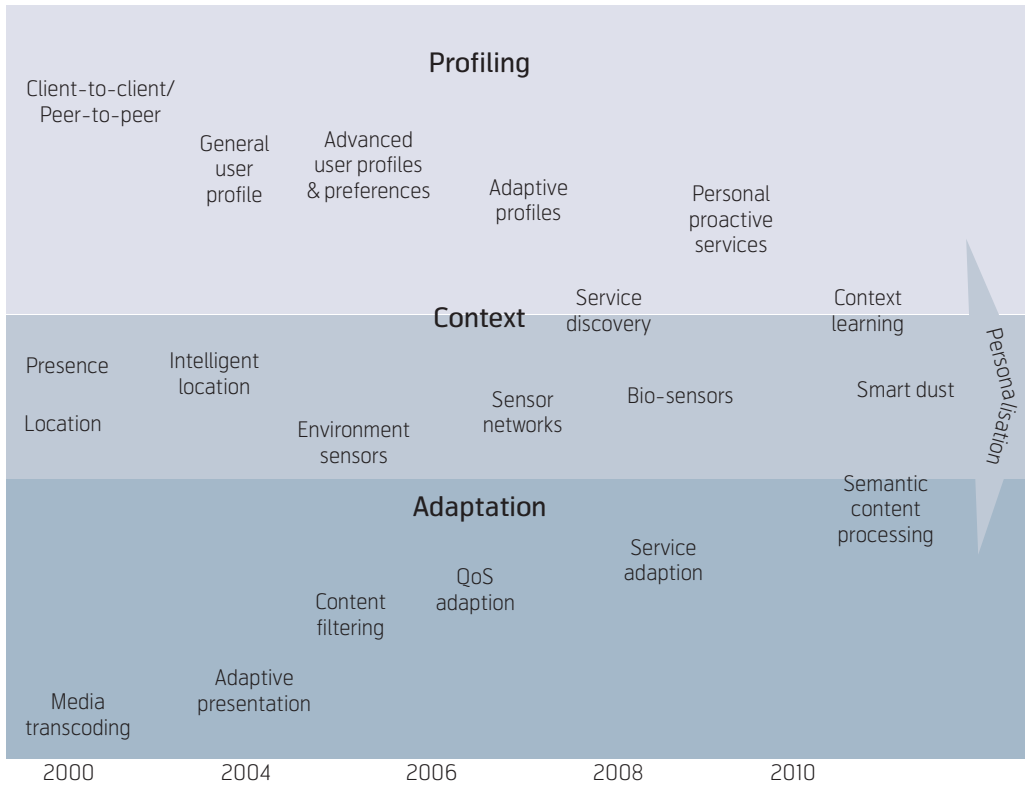


Figure 5 The vision of personalised services [4]

2 Personal Networks

Much of this material found in this section originates from the work done in the MAGNET project [6].

2.1 Introduction to Personal Networks

Having devices within a short range, and having these communicating devices to hold a personal relation to each other can be perceived as a Personal Area Net-

work (PAN). This concept has now existed for some time, and considering the development towards 4G, a natural extension to the PAN concept would be a Personal Network. Such a network is considered as a Personal Network, and is first described in [7] and [8]. A Personal Network (PN) is a network that connects the user's PAN to remote networks, like other PANs, infrastructure networks in buildings, or home

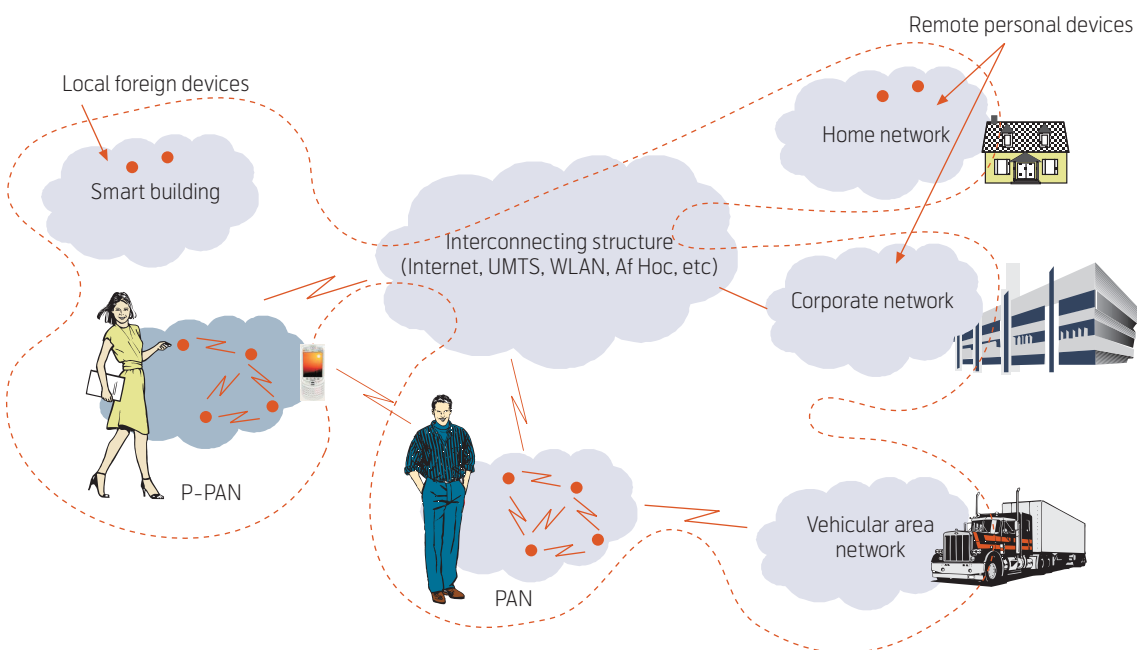


Figure 6 Conceptual illustration of Personal Networks [6]

networks. Figure 6 illustrates the concept of PNs, showing its heterogeneous collection of networks.

In a Personal Network, a user is able to connect to his/her devices and services using whatever infrastructure is available for communication. Personal networks are dynamic in the sense that they are created, maintained and destructed in an ad hoc manner, e.g. when a user moves around a building, nodes become a part of the network ad hoc and may also leave the network as they come out of range or for other reasons no longer are useful to the user. In fact, a Personal Network can be defined as:

A dynamic overlay network of interconnected local and remote personal devices organized in clusters, which are connected to each other via some interconnecting structure.

An interconnecting structure includes: Internet, intranet, WLAN, UMTS, GSM, PSTN, ad hoc network etc.

For Personal Networks the communication between clusters is characterised by security mechanisms that ensure privacy and protect the devices within the network from outside attacks. This means that security is a key issue in Personal Networks, since sensitive information regarding the user and private services will be accessible to the user anytime, anywhere,

while required to be protected against intruders. In a few words; Personal Networks can be characterised by:

- Trusted communication between personal device through pre-established trust relations between personal devices;
- Automatic formation in ad hoc fashion (no manual configuration or authentication steps);
- Support of professional and private services.

The introduction of Personal Networks is really what captures the convergence of communication technology at all layers for 4G which is highly desirable, and in this context Personal Networks are a very potent option for future communication networks.

2.2 Building blocks of Personal Networks

A Personal Network is constructed of several components. In Table 1, an overview can be found of the most important components and terminology used [9].

The composition of these entities and components is what makes the Personal Network and leads to a definition of the PN at a network level. However, the network level is not the only part of the PN. One of the key requirements to the PN is that it can provide access to services at any time, anywhere, which means that the heterogeneity of access technology must be taken into account as well.

2.3 Federation of Personal Networks

In a person's everyday situation he/she will not only need to interact with his or her applications and services. In fact, most persons have a family to interact with, a job, friends or other relations with whom they could eventually need to share information, resources, services or applications. In order to ensure a maximum utilisation of PNs, the components in a PN must support that a user wishes to also share services, resources and information with his/her friends, family, co-workers and so on. In these settings, PN needs to be federated between the involved groups which the user wishes to interact with. The interaction can then take on various roles in the communication, e.g. a person may wish to share more and secret information with his wife, than with a co-worker and so on. Figure 7 illustrates the basic principle of federated networks.

For federated PNs security and privacy become even more important issue than to PNs, since now users will be able to interact and share services, resources, information and so on, which are sensitive to malicious persons. Trust relations and authentication of

Personal Node	A node related to a given user or person with a pre-established trust attribute
Foreign Node	A node that is not personal and is not a part of the PN
Personal Device	A device related to a given user or person with a pre-established trust attribute
Foreign Device	A device that is not personal and not a part of the PN
Personal Service	Personal services are provided by personal nodes and devices and are available only to personal nodes and devices
Public Service	Public services can be given by any device/node (both personal and foreign)
Private Personal Area Network	A Private Personal Area Network or P-PAN is a dynamic collection of personal nodes and devices around a person
Cluster	A network of personal devices and nodes located within a limited geographical area (such as a house or a car) which are connected to each other by one or more network technologies and characterised by a common trust relationship between each other
Personal Network	A Personal Network (PN) includes the P-PAN and a dynamic collection of remote personal nodes and devices in clusters that are connected to each other via Interconnecting Structures

Table 1 Components and entities used to construct a Personal Network [9]

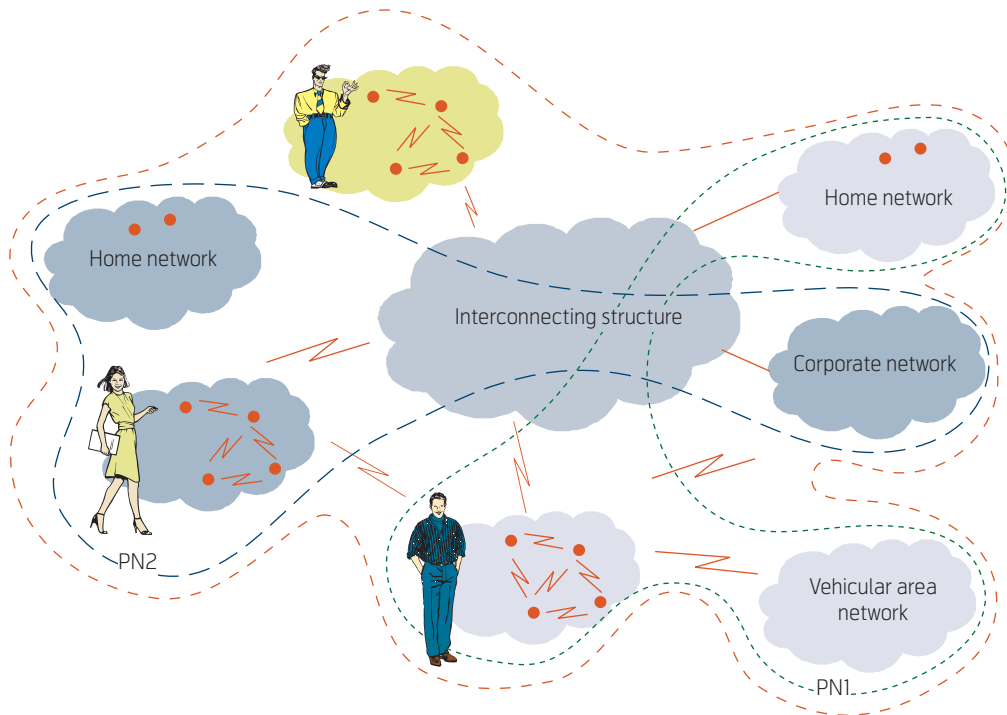


Figure 7 Federation of Personal Networks [10]

invited users will become an even more important issue since the possible combinations of interaction between users in the world are extremely high, and not all users one interacts with may be as nice as we would like to believe. The IST project MAGNET Beyond [11] takes up the challenge of moving Personal Networks this step further, towards federation of Personal Networks.

3 Abstraction levels of Personal Networks

A Personal Network can be seen from various levels, which is illustrated in Figure 8.

The differentiation of the Personal Network into the various levels seen in the figure, is to accommodate the technical challenges found at each level. The cooperation between the levels is very important, as this will ultimately enable the convergence of the technologies and ultimately lead the way towards 4G. In the following sub sections, a brief description of what technical challenges are addressed at each level is given.

3.1 Service abstraction level

At the service abstraction level, service, resource and context discovery and management are key issues. Such systems enable users to automatically discover services, whether public or private, that are within connectivity range of the user's device. Context discovery is used to discover where to obtain context information which later can be used to achieve con-

text awareness, i.e. having applications, services or other components in the PN architecture to change their behaviour and adapt to the given contextual situation a user might be in. This section provides a high level view of the architecture and components used to achieve this. For details on these topics see [12] and [13].

3.2 Network abstraction level

A PN is defined at the network level and is constructed by the components described in Table 1. The main objectives and challenges for this level are the establishment and interaction with the P-PAN and the remote clusters through the interconnecting structure. Some of the key elements for this to happen are [6]:

- Naming, addressing and intra PN routing
- The ad hoc self configuration for PN establishment
- Tunnelling mechanisms
- Mobility management and support issues

However, as can be seen in Figure 8 the infrastructure is an indispensable part of a PN for connecting to the remote nodes. The infrastructure, which may also include access networks, is expected to be a heterogeneous network consisting of various types of wired, wireless, public, private and shared networks. However, there are challenges concerning the communication between a cluster and the infrastructure that must be handled in the connectivity, as well as in the network layer. Such challenges are, for example, access network detection, access network announcement, selection of most suitable access technology taking

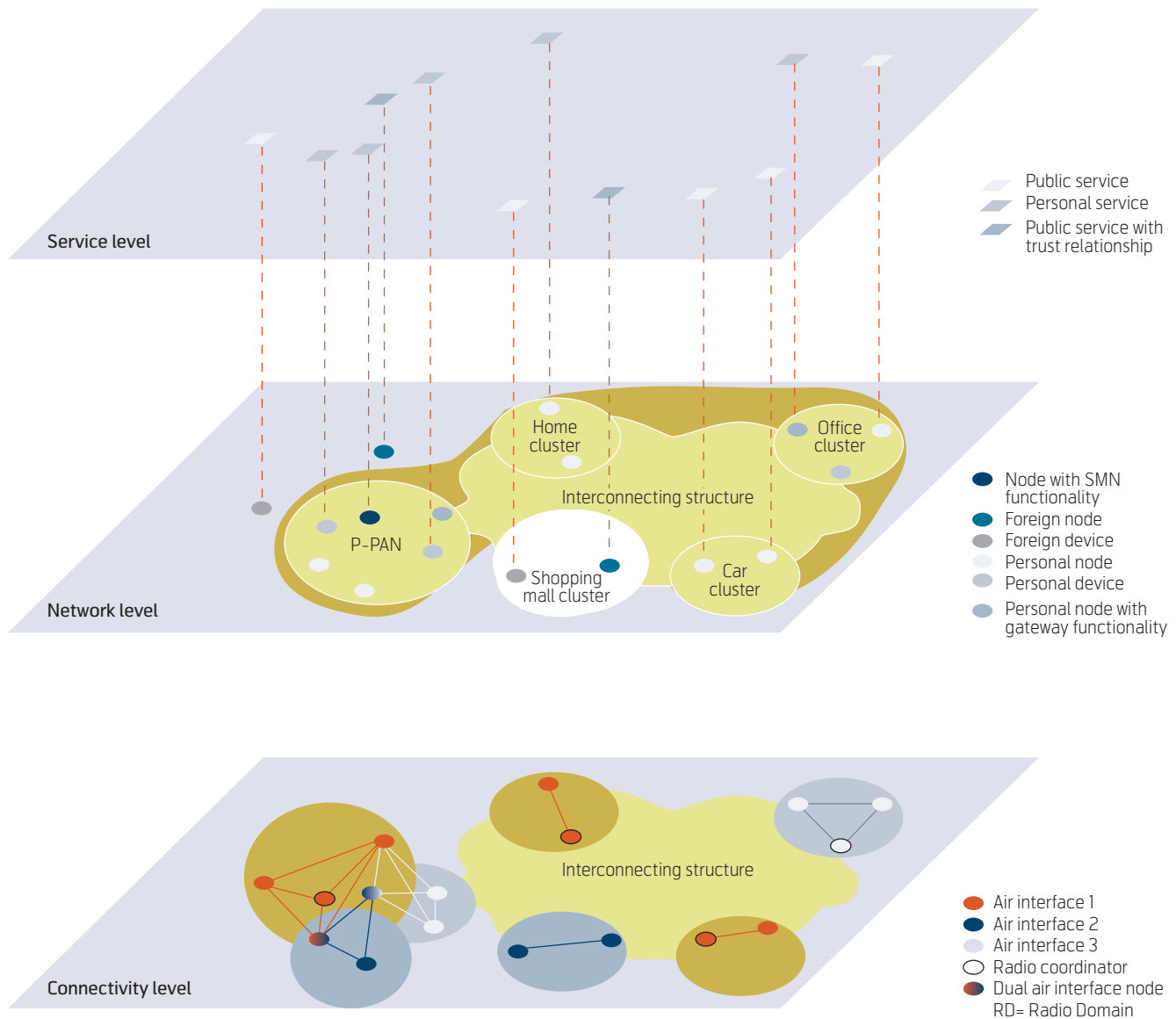


Figure 8 Abstraction levels in Personal Networks [9]

into account e.g. context information (user, environment, network), multiple gateway nodes management, traffic scheduling, access network monitoring and control, Radio Resource Management and fast handover schemes. More information on these issues can be found in [14].

3.3 The connectivity abstraction level

The connectivity abstraction level as shown in Figure 8 is composed of various radio domains corresponding to a given radio technology coordinated by a single radio coordinator. A radio domain describes an area within which devices can communicate with each other up to the MAC level using a common MAC control channel. This control channel is used, amongst other tasks, to evaluate in a coordinated fashion if a data channel can be used for communications between two or more devices. Within each radio domain, the objective is to minimise interference and optimise the amount of watts required for each bit

transmitted within one radio technology in order to offer PN users the best available radio link quality according to surrounding propagation and interference conditions [9].

4 Security and privacy in Personal Networks

Current security solutions for wireless technologies such as the one from 3GPP for GSM/UMTS based on (U)SIM algorithms, IEEE 802.11i drafts for WLAN security, or the Bluetooth security recommendations, are all tailored to securing the traffic exchanged between user devices and access points. While combining both transport and application layers, which would result in end-to-end secure communication, this approach was optimized for the Internet with a fixed infrastructure and to powerful devices capable of decrypting the secure message. To achieve the same level of security in ad-hoc networks, app-

roaches based on co-operative authorization and distributed key management are to be considered. However, these solutions usually incur considerable overheads in terms of signalling, and thus bandwidth usage, and processing needs. To make PN happen, security and privacy must be taken seriously. Trying to cover all aspects of security in one go is not possible, but in the following sections we discuss some important security aspects.

4.1 Threats and attacks

Personal Networks are subject to containing sensitive information about the user, and also services which the user may not wish to share with others, in particular with malicious users. There exist many parties in the world that wish a user no good; people wishing to trick credit card information from a user with fake services, terrorists wishing to destroy infrastructure, hackers just wanting to show off, etc., etc. Table 2 shows just a subset of the possible attacks that are potentially threatening Personal Networks [15].

One thing is for sure; security must be applied in all layers to overcome all the possible attacks that may occur, including encryption in PHY/MAC, authentication, authorisation, key establishment and secure communication.

At the highest level, the objective is to create an environment where message level transactions and business processes can be conducted securely in an end-to-end fashion. There is a need to ensure that messages are secured during transit, with or without the presence of intermediaries. There may also be a need to ensure the security of the data in storage. The following requirements are a few examples for providing end-to-end security for the PN/P-PAN [15]

- Authentication Mechanisms
- Authorisation and Access Control Mechanisms
- Confidentiality and Integrity
- Availability, User privacy and Non-Repudiation.

4.2 Trust establishment and management

The establishment of trust relationship between nodes and devices in the P-PAN and PN is a crucial aspect in the establishment of a Personal Network, simply because all services and applications are depending on this trust. If nodes are not trusted, information and services residing on the node cannot be trusted either, and may compromise the security level in the whole PN. Therefore, the establishment of trust relation between nodes is of utmost importance.

Because of the ad-hoc nature of the PN networks it is important for security mechanisms to be able to work in such a dynamically changing environment, and not

necessarily have access to a global network. Therefore a decentralised trust management seems to be the most appropriate. In such a management model all the participating devices will have the necessary information to trust other devices and create a virtual network between them.

However, in general one cannot assume that only two nodes are connected without a third one intercepting their communication (hence a potential man in the middle attacker). In order to create a management system, there is a need for appropriate certificates signed from authenticated users and a mechanism to share them and be able to validate and trust them online. Furthermore, the trust establishment and management system for PNs must be robust and flexible in order for the PN to become productive.

5 Personal Networks – a user adaptive network

Personal Networks is to a high degree also about Adaptability, i.e. the ability to change behaviour to the most appropriate under the given circumstances. It is very important that such an ability happens automatically, since the user has no interest in reconfiguration of a vast number of parameters every time some change happens in either the physical world or in the network. Adaptability can happen at several

Denial of Service (DoS)	Denial of service (DoS) attacks focus on preventing legitimate users of a service from the ability to use the service.
Man-in-the-Middle attack	Could be done by faking a service identity in order to make users send sensitive data to it.
Message alteration	In this attack type an attacker may modify parts or the whole message (including header and body parts), or delete part of that, or even insert some extra information into the message to achieve whatever goal the attacker has set
Eavesdropping	In this threat, unauthorised entities obtain access to information within a message or message parts.
Spoofing	In this type of attack, the attacker assumes the identity of a trusted entity in order to sabotage the security of the target entity. As far as the target entity knows, it is carrying on a conversation with a trusted entity.
Replay attack	In this attack an intruder intercepts a message and then replays it back to a targeted agent.
Phishing attack	Phishing attacks use spoofed e-mails and fraudulent Web sites to lure users into entering personal data such as credit card numbers, account usernames and passwords, which can then be used for financial theft or identity theft.

Table 2 A list of potential threats toward Personal Networks

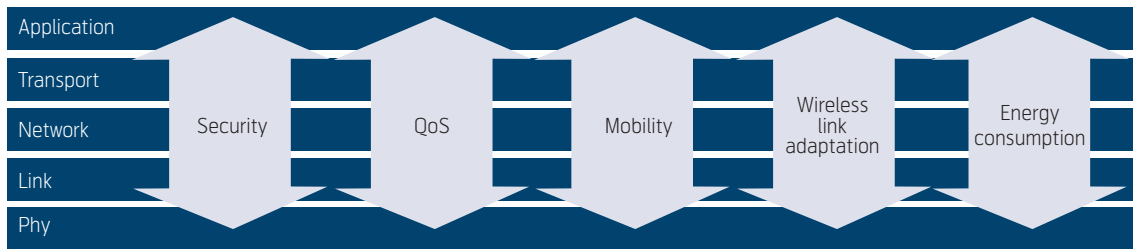


Figure 9 Cross layer optimisation issues

levels using different information. In the following, we describe three different types of adaptability that are under research and development.

5.1 Adaptive and Scalable Air Interfaces for WPANs

Traditional layered network architecture is very useful, but turns out to be too inflexible for the wireless domain. Figure 9 shows how information found on different network layers may flow up and down in order to enhance performance of a specific layer.

Allowing information to flow vertically in this way has shown many advantages and is still under research, while the price of this is an increased complexity in the software development. Knowing what application is running and needs to communicate, may help the physical link or other layer to choose what parameter to use, what air interface to use, etc. The other way round; information on the current conditions on e.g. link level or network level can also help the application to make certain choices, and thereby adapt its behaviour to the current situation.

5.2 Context aware service discovery

At the service level, service discovery is used for discovery services within the network. However, just searching for services is not enough for PN. Just the size of devices and nodes within a PN alone gives an indication of how many services potentially will be in a PN for the user. How can a user figure out which is the most relevant service under the given circumstances. Having service discovery to be context aware enables this system to react and decide for the user, based on the context of a user, which service might be the appropriate one to use. A simple example is the discovery of a nearby printer, i.e. the system knows the current location of a user, and the user needs a printer nearby which is not occupied by someone, i.e. is not busy printing a huge document. A context aware service discovery system will allow finding the nearest printer, which is also available, and otherwise meet requirements set by the user and his/her preferences.

5.3 Adaptive security

Determining what kind of security means are needed is very tightly coupled to the weight between required security level, time and energy that are required for a certain algorithm or protocol to do their job. A user does not want to fiddle and change settings each time a new application is started and needs to communicate. Therefore, a security architecture is required which is aware of whatever service or application the user wishes to use. The adaptivity of security in Personal Networks also closely relates to the user profiles, i.e. who you are, what kind of work you have and so on. Local security policy also plays a major role when considering adaptive security.

6 Conclusions

In this paper we discussed and showed that the progression towards 4G communication network is no longer a question of a linear development of higher bandwidth, less delay etc., but is also about convergence of technology. We defined in this paper 4G as an evolutionary and revolutionary new fully IP-based integrated system of systems and network of networks achieved after convergence of wired and wireless networks as well as computers, consumer electronics, and communication technology and several other convergences that will be capable to provide 100 Mb/s and 1 Gb/s, respectively in outdoor and indoor environments, with demand-driven end-to-end QoS and high security, offering any kind of services at any time as per user requirements, anywhere with seamless interoperability, always on, at an affordable cost, with one billing and fully personalized.

Personal Networks aims to reach these goals, making specifically this network paradigm an important step towards 4G. The fact that Personal Networks are strongly personal and are centred around the user, are key features for next generation networks. Looking at the composition of a Personal Network, it is clearly seen that such a network will operate on levels ranging from PHY to Application layer, and the development of such a network paradigm will address the issues pursued in 4G. Beside convergence, security

and privacy is an important part as the user will have to trust these networks with sensitive information, and the communication that happens at all layers. Study and development of these issues as well as new business models for PNs and PN specific applications are currently carried out in the IST project MAGNET Beyond [11] as a continuation of the MAGNET project [6]. To reflect on our title of this paper, *The unpredictable future: Personal Networks paving towards 4G*, it is clear that nobody knows the future, but that the realisation of Personal Networks for sure will be an important step towards 4G. In our view, 4G can be defined by the following equation [16]:

$$B3G + Pers \triangleq 4G$$

where *B3G* stands for beyond third generation, which is defined as the integration of existing systems to interwork with each other and with the new interface. *Pers* stands for personalisation, which is the key issue in Personal Networks.

References

- 1 Kim, Y K, Prasad, R. *4G Roadmap and Emerging Communication Technologies*. Artech House, 2006.
- 2 Bria, A, Gessler, F, Queseth, O, Stridh, R, Unbehau, M, Wu, J, Zander, J. 4th-Generation Wireless Infrastructures: scenarios and Research Challenges. *IEEE Personal Communications*, 25–31, December 2001.
- 3 Prasad, A R, Zugenmaier, A, Schoo, P. Next generation Communications and Secure Seamless Handover. *First IEEE/CREATE-Net Workshop on Security and QoS in Communication Networks*, Athens, Greece, 2005
- 4 *PRODEMIS Project: Global Technology Roadmap*. Deliverable, August 2004. (IST-2000-26459) URL: <http://www.prodemis-ist.org>
- 5 Ghader, M, Olsen, R L, Prasad, N, Mirzadeh, S, Tafazolli, R. *Secure Resource and Service Discovery in Personal Networks*. WWRF12 meeting, Toronto, 2005.
- 6 *My personal Adaptive Global NET (MAGNET)*. (IST 507102) URL: <http://www.ist-magnet.org>
- 7 Niemegeers, I G, de Groot, S M H. Personal Networks: Ad Hoc Distributed Personal Environments. *Med-HocNet, IFIP Conference on Ad-Hoc Networks*, September 2002.
- 8 Niemegeers, I G, de Groot, S M H. From Personal Area Networks to Personal Networks: A User Oriented Approach. *Personal Wireless Communication*, Kluwer, May 2002.
- 9 Petrova, M et al. *Overall secure PN architecture*. IST 507102 MAGNET Deliverable 2.1.2, November 2005. Available at <http://www.ist-magnet.org>
- 10 Niemegeers, I G, de Groot, S M H. FEDNETS: Context-Aware Ad-Hoc Network Federations. *Wireless Personal Communication*, Springer, 33, June 2005.
- 11 IST. *MAGNET Beyond*. (IST-FP6-IP-027396) URL: <http://www.magnet.aau.dk>
- 12 Ghader, M et al. *Resource and Service Discovery: PN Solutions*. IST 507102 MAGNET Deliverable 2.2.1, December 2004. Available at <http://www.ist-magnet.org>
- 13 Olsen, R L et al. *Service, Resource and Context Discovery system specification*. MAGNET Deliverable D2.2.3, December 2005. To be publicly available at <http://www.ist-magnet.org>
- 14 Jacobsen, M et al. *Refined Architectures and Protocols for PN Ad-hoc Self-configuration, Interworking, Routing and Mobility Management*. MAGNET Deliverable D2.4.3, December 2005. To be publicly available at <http://www.ist-magnet.org>
- 15 Rebahi, Y et al. *State of the art and Functional Specification of Service-Level Security Architecture*. MAGNET Deliverable D4.2.1, September 2004. Available at <http://www.ist-magnet.org>
- 16 Prasad, R, Deneire, L. *From WPANs to Personal Networks – Technologies and Applications*. Artech House, 2006. (ISBN-10:1-58053-826-6)

Ramjee Prasad was born in Babhaur (Gaya), Bihar, India, on July 1, 1946. He is now a Dutch Citizen. He received a B.Sc. (Eng) degree from Bihar Institute of Technology, Sindri, India and M.Sc. (Eng) and Ph.D. degrees from Birla Institute of Technology (BIT), Ranchi, India, in 1968, 1970 and 1979, respectively. Since June 1999, Dr. Prasad has been with Aalborg University, where he is currently Director of Center for Tele-infrastruktur (CTIF), and holds the chair of wireless information and multimedia communications. He is co-ordinator of European Commission Sixth Framework Integrated Project MAGNET (My personal Adaptive Global NET). He was involved in the European ACTS project FRAMES (Future Radio Wideband Multiple Access Systems) as a DUT project leader. He is a project leader of several international, industrially funded projects. He has published over 500 technical papers, contributed to several books, and has authored, co-authored, and edited sixteen books. He has served as a member of advisory and program committees of several IEEE international conferences. In addition, Dr. Prasad is the co-ordinating editor and editor-in-chief of the Springer International Journal on Wireless Personal Communications and a member of the editorial board of other international journals. Dr. Prasad is also the founding chairman of the European Center of Excellence in Telecommunications, known as HERMES, and he is now Honorary Chair. He has received several international awards; the latest being the "Telenor Nordic 2005 Research Prize" (website: <http://www.telenor.no/om/>). He is a fellow of IEE, a fellow of IETE, a senior member of IEEE, a member of The Netherlands Electronics and Radio Society (NERG), and a member of IDA (Engineering Society in Denmark). Dr. Prasad is advisor to several multinational companies.

email: prasad@kom.aau.dk

Rasmus Løvenstein Olsen was born in Aarhus, Denmark on June 12, 1977. He received his M.Sc. in Electrical Engineering from Aalborg University in 2003 with focus on antenna control for satellite communication. He has during his time at the university worked with the ASAP and PRODEMIS web pages. He has also been very active in the pico-satellite program of Aalborg University called AAU-Cubesat. In 2004 he started in the IST project MAGNET; where he has been working with Context Aware Service Discovery for Personal Networks. He is currently working toward his PhD degree in that area.

email: rlo@kom.aau.dk

Terms and acronyms in Real-time communication over IP

2G	Second Generation (mobile system)	Refers to the family of digital cellular telephone systems standardised in the 1980s and introduced in the 1990s. They introduced digital technology and carry both voice and data conversation. CDMA, TDMA and GSM are examples of 2G mobile networks.
3D	Three-dimensional	Something having three dimensions, e.g. width, length, and depth. A three-dimensional space, which appears to exist as three dimensions. A vector space or coordinate space with three dimensions.
3G	Third Generation (mobile system)	The generic term for the next generation of wireless mobile communications networks supporting enhanced services like multimedia and video. Most commonly, 3G networks are discussed as graceful enhancements of 2G cellular standards, like e.g. GSM. The enhancements include larger bandwidth, more sophisticated compression techniques, and the inclusion of in-building systems. 3G networks will carry data at 144 kb/s, or up to 2Mb/s from fixed locations. 3G will standardize mutually incompatible standards: UMTS FDD and TDD, CDMA2000, TD-CDMA.
3GPP	Third Generation Partnership Project	Group of the standards bodies ARIB and TTC (Japan), CCSA (People's Republic of China), ETSI (Europe), T1 (USA) and TTA (Korea). Established in 1999 with the aim of producing the specifications for a third generation mobile communications system called UMTS. A permanent project support group called the Mobile Competence Centre (MCC) is in charge of the day to day running of 3GPP. The MCC is based at the ETSI headquarters in Sophia Antipolis, France. http://www.3gpp.org
AAA	Authentication, Authorization and Accounting	Key functions to intelligently controlling access, enforcing policies, auditing usage, and providing the information necessary to do billing for services available on the Internet.
AAC	Advanced Audio Coding	Audio Coding algorithm defined in MPEG 2 (ISO/IEC Publication 13818) and MPEG 4 (ISO/IEC Publication 14496) standards. http://www.iso.org
AAC-LD	Low Delay AAC	Low delay version of AAC defined in MPEG 4 (ISO/IEC Publication 14496). http://www.iso.org
AAL	ATM Adaptation Layer	The use of Asynchronous Transfer Mode (ATM) technology and services creates the need for an adaptation layer in order to support information transfer protocols, which are not based on ATM. This adaptation layer defines how to segment and reassemble higher-layer packets into ATM cells, and how to handle various transmission aspects in the ATM layer. Examples of services that need adaptations are Gigabit Ethernet, IP, Frame Relay, SONET/SDH, UMT/Wireless, etc. The main services provided by AAL are: Segmentation and reassembly, handling of transmission errors, handling of lost and misinserted cell conditions and timing and flow control. http://www.itu.int
A/D	Analogue to Digital	The process of converting continuous signals to discrete digital numbers.
AbS	Analysis-by-Synthesis	SA speech coding principle.
AC	Access Category	
ACELP	Algebraic CELP	A speech coding algorithm.
ACF	Admission Confirm	
ACK	Acknowledgement	A packet used in e.g. TCP to acknowledge receipt of a packet.
ACR	Absolute Category Rating	
ADPCM	Adaptive Differential PCM	A speech coding algorithm where the difference between the present sample and the previous one is encoded using adaptive quantization intervals.
ADSL	Asymmetric Digital Subscriber Line	A data communications technology that enables faster data transmission over copper telephone lines than a conventional modem can provide. The access utilises the 1.1 MHz band and has the possibility to offer, depending on subscriber line length, downstream rates of up to 8 Mb/s. Upstream rates start at 64 kb/s and typically reach 256 kb/s but can go as high as 768 kb/s.

AES	Advanced Encryption Standard	It is also known as Rijndael. In cryptography, it is a block cipher adopted as an encryption standard by the US government. It is expected to be used worldwide and is analysed extensively, as was the case with its predecessor, the Data Encryption Standard (DES). AES was adopted by National Institute of Standards and Technology (NIST) in November 2001 after a 5-year standardisation process. The cipher was developed by two Belgian cryptographers, Joan Daemen and Vincent Rijmen, and submitted to the AES selection process under the name "Rijndael", a blend comprising the names of the inventors. www.nist.gov
AF	Assured Forwarding	
AIB	Authenticated Identity Body	
AIFS	Arbitrary Interframe Space	
AIFSN	Arbitrary Interframe Space Number	
AIP	All IP	
AIPN	All IP Network	
ALG	Application Level Gateway	
A-MGF	Access Media Gateway Function	
AMR	Adaptive Multi Rate	A speech coding algorithm offering a wide range of data rates. Designed for use in 3G networks. The philosophy behind AMR is to lower the codec rate as the interference increases and thus enabling more error correction to be applied. The 12.2 kbit/s mode is equivalent to the GSM Enhanced Full Rate (EFR) codec.
AMR-WB	Adaptive Multi Rate Wideband	A speech coding algorithm offering a wide range of data rates. Designed for use in 3G networks. The codec is also defined in ITU-T Recommendation G.722.2. http://www.3gpp.org , http://www.itu.int
AOL	America Online	A US-based online service provider, Internet service provider, and media company operated by Time Warner. Based in Dulles, Virginia, with regional branches around the world, it is by far the most successful proprietary online service, with more than 32 million subscribers at one point in the US, Canada, Germany, France, the United Kingdom, Latin America (declared bankrupt in 2004), Japan and formerly Russia. In early 2005, AOL Hong Kong stopped its service. In the fall of 2004, AOL reported total subscribers had dropped to 24 million, a drop of over a quarter of its subscribers. In late 1996, AOL suspended all dialup service within Russia in the face of massive billing fraud, forcing the company into a rare case of full market retreat. http://www.aol.com/
AP	Access Point	A point where users access the system/network, e.g. a base station in a wireless network.
API	Application Programming Interface	The specific method prescribed by a computer operating system or by an application program by which a programmer writing an application program can make requests of the operating system or another application. A set of routines, protocols, and tools for building software applications. Most operating environments, such as MS-Windows, provide an API so that programmers can write applications consistent with the operating environment.
ARJ	Admission Reject	Message for signalling between end-point and Gatekeeper specified in ITU-T Recommendation H.225.0. http://www.itu.int
ARQ	Admission Request	Message for signalling between end-point and Gatekeeper specified in ITU-T Recommendation H.225.0. http://www.itu.int
ARQ	Automatic Repeat reQuest	A standard method of checking transmitted data used on high-speed data communications systems. The sender encodes an error-detection field based on the contents of the message. The receiver recalculates the check field and compares it with the received one. If they match an 'ACK' (acknowledgement) is transmitted to the sender. If they do not match, a 'NAK' (negative acknowledgement) is returned, and the sender retransmits the message.
AS	Application Server	

ASCII	American Standard Code for Information Interchange	A character encoding based on the English alphabet. ASCII codes represent text in computers, communications equipment, and other devices that work with text. Most modern character encodings have a historical basis in ASCII. It was first published as a standard in 1967 and was last updated in 1986. It currently defines codes for 33 non-printing, mostly obsolete control characters that affect how text is processed, plus 95 printable characters. ASCII was subsequently updated and published as ANSI X3.4-1968, ANSI X3.4-1977, and finally, ANSI X3.4-1986. http://www.ansi.org
ASN.1	Abstract Syntax Notation One	A language used by the OSI protocols for describing abstract syntax.
AT&T	The American Telephone and Telegraph Company	AT&T Inc. is based in San Antonio, Texas. It is the largest provider of both local and long distance telephone services and wireless service in the United States. Originally founded in 1885 as the American Telephone and Telegraph Company. The modern company was formed in 2005 by Southwestern Bell Corporation's (SBC Communications) purchase of its former parent company, AT&T Corp.
ATA	Analogue Telephone Adapter	An analogue telephone adapter is a device used to connect an analogue telephone to a computer or network so that the user can make calls over the Internet or other digital networks.
ATM	Asynchronous Transfer Mode	A high bandwidth, low-delay, connection-oriented, packet-like switching and multiplexing technique. ATM allocates bandwidth on demand, making it suitable for high-speed connections of voice, data and video services. Access speeds are up to 622 Mb/s and backbone networks currently operate at speeds as high as 2.5 Gb/s. Standardised by ITU-T [Newton03].
AUC	Authentication Centre	
AVC	Advanced Video Coding	Often used to identify the most recent MPEG/ITU-T video coding algorithm (ITU-T Recommendation H.264). http://www.itu.int
B2BUA	Back to Back User Agent	The Back-To-Back User Agent is a Session Initiation Protocol (SIP) based logical entity that can receive and process INVITE messages as a SIP User Agent Server (UAS). It also acts as a SIP User Agent Client (UAC) that determines how the request should be answered and how to initiate outbound calls. Unlike a SIP proxy server, the B2BUA maintains complete call state and participates in all call requests. http://www.whatis.com
BB	Bandwidth Broker	
BBT	Broadband Telephony	Broadband Telephony is the utilisation of broadband connections to deliver voice calls. Calls are transmitted as IP Packets to the host company, where they either 'break out' to the public networks, or continue as IP calls across the Internet. Usually synonymous to "VoIP telephony" or "Internet telephony".
BE	Best Effort	Used to identify an IP network service where no priority is given to the traffic.
BGCF	Breakout Gateway Control Function	
BGF	Border Gateway Function	
BGWF	Border Gateway Function	
BICC	Bearer Independent Call Control	
BK	Background	
BO	Backoff	
BRI	Basic Rate Interface	An ISDN user-network access arrangement that corresponds to the interface structure composed of two B-channels and one D-channel. The bit rate of the D-channel for this type of access is 16 kbit/s.
BSS	Basic Service Set	A Basic Service Set (BSS) is the basic building block of an IEEE 802.11 wireless LAN (according to the IEEE802.11-1999 standard). The most basic BSS is two STAs in IBSS mode. In infrastructure mode, a basic BSS consists of at least one STA and one Access Point (AP). http://www.ieee802.org

BT	Bluetooth	Bluetooth is an industrial specification for wireless personal area networks (PANs). Bluetooth provides a way to connect and exchange information between devices like personal digital assistants (PDAs), mobile phones, laptops, PCs, printers and digital cameras via a secure, low-cost, globally available short range radio frequency. https://www.bluetooth.org/
CAMEL	Customised Applications for Mobile network Enhanced Logic	A set of GSM standards designed to work on a GSM core network. They allow an operator to define services over and above standard GSM services. The CAMEL architecture is based on the Intelligent Network (IN) standards, and uses the CAP protocol.
CAP	Controlled Access Periods	
CBC-MAC	Cipher Block Chaining-Message Authentication Code	
CCI	Controlled Contention Intervals	
CCITT	Comité Consultatif International Téléphonique et Télégraphique	The International Telegraph and Telephone Consultative Committee of the ITU, called ITU-T as from 1992. http://www.itu.int/
CCK	Complementary Code Keying	Complementary Code Keying (CCK) is a modulation scheme used with wireless networks (WLANs) that employ the IEEE 802.11b specification. A complementary code contains a pair of finite bit sequences of equal length, such that the number of pairs of identical elements (1 or 0) with any given separation in one sequence is equal to the number of pairs of unlike elements having the same separation in the other sequence.
CCMP	Counter mode with CBC-MAC Protocol	
CD	Compact Disc	An optical disc used to store digital data, originally developed for storing digital audio. A standard compact disc, often known as an audio CD to differentiate it from later variants, stores audio data in a format compliant with the Red Book standard. An audio CD consists of several stereo tracks stored using 16-bit PCM coding at a sampling rate of 44.1 kHz. Standard compact discs have a diameter of 120 mm, though 80-mm versions exist in circular and "business-card" forms. The 120-mm discs can hold 74 minutes of audio, and versions holding 80 or even 90 minutes have been introduced. The 80-mm discs are used as "CD-singles" or novelty "business-card CDs". They hold about 20 minutes of audio. Compact disc technology was later adapted for use as a data storage device, known as a CD-ROM. The first edition of the Red Book was released in June 1980 by Philips and Sony; it was adopted by the Digital Audio Disc Committee and ratified as IEC 908. http://www.iec.ch
CDF	Cumulative Distribution Function	In probability theory, the Cumulative Distribution Function completely describes the probability distribution of a real-valued random variable, X. It represents the probability that the random variable X takes on a value less than or equal to x.
CELP	Code-Excited Linear Predictive	A speech coding algorithm.
CFB	Contention Free Bursting	
CFP	Contention Free Period	
CL	Controlled Load	
CLIP	Calling Line Identification Presentation	A telephony intelligent network service that transmits the caller's telephone number to the called party's telephone equipment during the ringing signal or when the call is being set up but before the call is answered.
CLIR	Calling Line Identification Restriction	Contrary to CLIP, blocking the transmission and display of the caller's number at the called party's telephone equipment.
CN	Core Network	
CNR	Carrier-to-Noise Ratio	
COLP	Connected Line Presentation	A supplementary service.
COLR	Connected Line Restriction	A supplementary service.

COPS	Common Open Policy Service	Part of the internet protocol (IP) suite as defined by the IETF RFC 2748. It specifies a simple client/server model for supporting policy control over Quality of Service (QoS) signalling protocols (e.g. RSVP). Policies are stored on servers, also known as Policy Decision Points (PDP), and are enforced on clients, also known as Policy Enforcement Points (PEP). http://www.ietf.org
CP	Contention Period	A time period when two or more data stations attempt to transmit at the same time over a shared channel, or when two data stations attempt to transmit at the same time in two-way alternate communication.
CRM	Customer Relationship Management	An integrated information system that is used to plan, schedule and control the presales and post sales activities in an organisation. The objective is to enable a customer to interact with a company through various means including the web, telephone, fax, e-mail, mail and receive a consistent level of quality service.
CS	Circuit Switched	A network that establishes a circuit (or channel) between nodes before they may communicate. This circuit is dedicated and cannot be used for other means until the circuit is cancelled/closed and a new one created. If no actual communication is taking place in this circuit then the channel remains idle.
CSB	Circuit Switched Bearer	
CSCF	Call/Session Control Function	Several roles of SIP servers or proxies used to process SIP signalling packets in the IP Multimedia Subsystem (IMS).
CSMA/CA	Carrier Sense Multiple Access with Collision Avoidance	A network control protocol in which a carrier sensing scheme is used. A data station that intends to transmit sends a busy signal. After waiting a sufficient time for all stations to receive the busy signal, the data station transmits a frame. While transmitting, if the data station detects a busy signal from another station, it stops transmitting for a random time and then tries again. CSMA/CA is a modification of pure Carrier Sense Multiple Access (CSMA). Collision avoidance is used to improve the performance of CSMA by attempting to reserve the network for a single transmitter. This is the function of the "busy signal" in CSMA/CA. The performance improvement is achieved by reducing the probability of collision and retry. Extra overhead is added due to the busy signal wait time, so other techniques give better performance. Collision avoidance is particularly useful in media such as radio, where reliable collision detection is not possible. Apple's LocalTalk implemented CSMA/CA on an electrical bus using a three-byte busy signal. 802.11 RTS/CTS implements CSMA/CA using short Request to Send and Clear to Send messages.
CSN	Circuit-Switched Networks	See CS.
CTS	Clear To Send	
CW	Contention Window	
D/A	Digital to Analogue	The process of converting a digital (usually binary) code to an analogue signal (current, voltage or charges).
DCF	Distributed Coordination Function	A type of Medium Access Control (MAC) technique used in Wi-Fi Wireless LANs. DCF manages the transmission over a medium by allowing each node to listen to surrounding nodes to see if they are transmitting, before transmitting themselves. DCF is de-facto default setting for Wi-Fi hardware. http://www.ieee802.org/11
DCR	Degradation Category Rating	
DCT	Discrete Cosine Transform	A Fourier-related transform similar to the discrete Fourier transform (DFT), but using only real numbers. It is equivalent to a DFT of roughly twice the length, operating on real data with even symmetry (the Fourier transform of a real and even function is real and even), where in some variants the input and/or output data are shifted by half a sample.
DDI	Direct-Dialling-In	A feature offered by telephone companies for use with their customers' PBX system, whereby the telephone company allocates a range of numbers all connected to their customer's PBX.
DECT	Digital Enhanced Cordless Telecommunication	Formerly called Digital European Cordless Telephone. An ETSI standard for digital portable phones, commonly used for domestic or corporate purposes. DECT is a cellular system with cell radii of 25 to 100 metres. DECT uses a net bit rate of 32 kbit/s. It operates in the frequency band from 1880 to 1900 MHz. The band is divided into 10 carriers, each with 2 x 12 timeslots. It can serve a traffic density of approx. 10,000 telephony channels per square kilometre. The DECT physical layer is a combined frequency division multiple access (FDMA) / time division multiple access (TDMA) system using time division duplex (TDD) to separate traffic in the two directions. http://www.etsi.org

DHCP	Dynamic Host Configuration Protocol	Dynamic Host Configuration Protocol (DHCP) is a client-server networking protocol. A DHCP server provides configuration parameters specific to the DHCP client host requesting, generally, information required by the client host to participate on an IP network. DHCP also provides a mechanism for allocation of IP addresses to client hosts. DHCP appeared as a standard protocol in October 1993. RFC 2131 provides the latest (March 1997) DHCP definition. The latest standard on a protocol describing DHCPv6, DHCP in an IPv6 environment, was published in July 2003 as RFC 3315. http://www.ietf.org
DiffServ	Differentiated Services	A method of trying to guarantee quality of service on large networks such as the Internet. http://www.ietf.org
DIFS	DCF Interframe Space	
DLP	Direct Link Protocol	
DNS	Domain Name System	A system that stores information associated with domain names in a Distributed Database on networks, such as the Internet. The domain name system (Domain Name Server) associates many types of information with domain names, but most importantly, it provides the IP address associated with the domain name. It also lists mail exchange servers accepting e-mail for each domain. In providing a worldwide keyword-based redirection service, DNS is an essential component of contemporary Internet use. Paul Mockapetris invented the DNS in 1983; the original specifications appear in IETF RFC 882 and 883. In 1987, the publication of RFC 1034 and RFC 1035 updated the DNS specification http://www.ietf.org
DoS	Denial of Service	An attack on a computer system or network that causes a loss of service to users, typically the loss of network connectivity and services by consuming the bandwidth of the victim network or overloading the computational resources of the victim system.
	Downstream	Identifies transmission from the network to the user equipment.
DPCM	Differential PCM	A speech coding algorithm where the difference between the present sample and the previous one is encoded.
DS	Distribution System	
DSCP	DiffServ Code Point	
DSL	Digital Subscriber Line	A family of technologies that provide a digital connection over the copper wires of the local telephone network. Its origin dates back to 1988, when an engineer at Bell Research Lab devised a way to carry a digital signal over the unused frequency spectrum. This allows an ordinary phone line to provide digital communication without blocking access to voice services. Bell's management, however, were not enthusiastic about it, as it was not as profitable as renting out a second line for those consumers who preferred to still have access to the phone when dialling out. This changed in the late 1990s when cable companies started marketing broadband Internet access. Realising that most consumers would prefer broadband Internet to a second dial out line, Bell companies rushed out the DSL technology that they had been sitting on for the past decade as an attempt to slow broadband Internet access uptake, to win market shares against the cable companies. As of 2004, DSL provides the principal competition to cable modems for providing high speed Internet access to home consumers in Europe and North America. The reach-restraints (line length from Central Office to Subscriber) reduce as data rates increase, with technologies like VDSL providing short-range links (typically "fibre to the curb" network scenarios). Example DSL technologies (sometimes called xDSL) include: ADSL (Asymmetric Digital Subscriber Line), HDSL (High Bit Rate Digital Subscriber Line), RADSL (Rate Adaptive Digital Subscriber Line), SDSL (Symmetric Digital Subscriber Line, a standardised version of HDSL), VDSL (Very high speed Digital Subscriber Line), G.SHDSL (ITU-T Standardised replacement for early proprietary SDSL).
DSLAM	Digital Subscriber Line Access Multiplexer	A DSLAM is a network device, usually at a telephone company central office, that receives signals from multiple customer Digital Subscriber Line (DSL) connections and puts the signals on a high-speed backbone line using multiplexing techniques. Depending on the product, DSLAM multiplexers connect DSL lines with some combination of asynchronous transfer mode (ATM), frame relay, or Internet Protocol networks.
DST	Destination	
DTMF	Dual Tone Multifrequency	An analogue inband access signalling system where the combination of two tones (frequencies) define each of the digits 0-9 and the symbols *, #, A, B, C and D.

DVD	Digital Versatile Disc	Formerly Digital Video Disc. Data storage format released in 1995. The discs have the same physical size as the CD, but the capacity is more than 7 times as high, approx. 4.7 GB on one side. The discs can have dual layers per side, thus a double-sided, dual-layer disc can store approx. 17 GB of data. Used for storing video, sound, computer software and data, games, etc. A single-sided, single-layer disc can store a typical feature film of 130 minutes with 8 different surround quality sound tracks. Available as read-only (DVD-Video, DVD-ROM), Once writable (DVD-R, DVD+R) and re-writable (DVD-RW, DVD+RW, DVD-RAM). http://www.dvdforum.org
ECN	Explicit Congestion Notification	A congestion avoidance scheme that uses marking packets instead of dropping them in the case of incipient congestion. The receivers of marked packets should return the information about marked packets to the senders, and the senders should decrease their transmit rate. To avoid heavy congestion, routers mark packets with probability depending on an average queue length. IETF RFC 3186. http://www.ietf.org
ECS	Electronic Communications Service	
EDA	Enterprise Digital Assistant	
EDCA	Enhanced Distributed Channel Access	
EDCF	Enhanced DCF	
EDGE	Enhanced Data for GSM Evolution	A modulation method for GSM and IS-136 TDMA networks, standardized by ETSI, that allows for wireless data transfer up to 384 kb/s. http://www.etsi.org , http://www.3gpp.org
EE	Excellent Effort	
EEA	European Economic Area	An agreement between the European Free Trade Association (EFTA) and the European Union (EU) from 1 January 1994. It was designed to allow EFTA countries to participate in the European Single Market without having to join the EU. The current members (contracting parties) are three of the four EFTA states – Iceland, Liechtenstein and Norway (not Switzerland) – the European Union and the 25 EU Member States.
EF	Expedited Forwarding	
EFR	Enhanced Full-Rate (speech coding)	Enhanced Full Rate or EFR or GSM-EFR is a speech coding standard that was developed in order to improve the quite poor quality of GSM-Full Rate (FR) codec. The EFR 12.2 kbit/s speech coding standard is compatible with the highest AMR mode. The speech quality is similar to that of the ITU-T Recommendation G.711 codec used in PSTN/ISDN.
EG	ETSI Guide	
EIA	Electronics Industries Alliance	A US trade organization that includes the full spectrum of US manufacturers. It is a partnership of electronic and high-tech associations and companies whose mission is promoting the market development and competitiveness of the US high-tech industry through domestic and international policy efforts. It is headquartered in Arlington, Virginia and comprises nearly 1,300 member companies whose products and services range from the smallest electronic components to the most complex systems used by defense, space and industry, including the full range of consumer electronic products. EIA is accredited by ANSI to help develop standards on electronic components, consumer electronics, electronic information, telecommunications, and Internet security. Called the Electronic Industries Association until 1997. http://www.eia.org
EIFS	Extended Interframe Space	
ENUM	TELEphone NUmber Mapping	
EP	ETSI Project	
EPP	ETSI Partnership Project	A standardisation project where two or more regional standards organisations are participating.

ERG	European Regulators Group	An independent body for reflection, debate and advice in the electronic communications regulatory field created by the European Commission Decision 2002/627/EC adopted on 29 July 2002. It is composed of the heads of the relevant national authorities, and acts as an interface between them and the European Commission in order to advise and assist the Commission in consolidating the internal market for electronic communications networks and services. http://erg.eu.int
ES	ETSI Standard	Telecommunication Standard (ETSI TS) developed by the European Telecommunications Standards Institute (ETSI). http://www.etsi.org
ESP	Encapsulating Security Payload	
ESPRIT	European Strategic Program on Research in Information Technology of the European Union	An integrated programme of industrial R&D projects and technology take-up measures. It is managed by DG III, the Directorate General for Industry of the European Commission. It was part of the EU's Fourth Framework Programme, which ran from 1994 to 1998. http://www.cordis.lu/esprit/
ESS	Extended Service Set	
ETSI	European Telecommunication Standards Institute	A non-profit membership organization founded in 1988. The aim is to produce telecommunications standards to be used throughout Europe. The efforts are coordinated with the ITU. Membership is open to any European organization proving an interest in promoting European standards. It was e.g. responsible for the making of the GSM standard. The headquarters are situated in Sophia Antipolis, France. http://www.etsi.org
EURES-COM	The European Institute for Research and Strategic Studies in Telecommunications	An organisation for collaborative R&D in telecommunications. Eurescom was founded in 1991 by major European network operators and service providers. Based in Heidelberg, Germany, the organisation provides services for initiating, managing and supporting distributed collaborative research programmes to network operators, service providers, suppliers and vendors who wish to collaborate on the issues facing the telecommunications industry. http://www.eurescom.de
E-UTRAN	Evolved UTRAN	Term used by the 3GPP for the next generation UMTS Terrestrial Radio Access Network. It is developed as part of 3GPP's Long Term Evolution (LTE) and System Architecture Evolution (SAE) work. It was initiated in 2004, and first release of specifications is expected finished in 2007. http://www.3gpp.org
FCC	Federal Communication Commission (USA)	An independent United States government agency, directly responsible to US Congress. The FCC was established by the Communications Act of 1934 and is charged with regulating interstate and international communications by radio, television, wire, satellite and cable. The FCC's jurisdiction covers the 50 states, the District of Columbia, and US possessions. The FCC is directed by five Commissioners appointed by the President and confirmed by the Senate for 5-year terms, except when filling an unexpired term. The President designates one of the Commissioners to serve as Chairperson. Only three Commissioners may be members of the same political party. None of them can have a financial interest in any Commission-related business. As the chief executive officer of the Commission, the Chairman delegates management and administrative responsibility to the Managing Director. The Commissioners supervise all FCC activities, delegating responsibilities to staff units and Bureaus. The Commission staff is organized by function. There are six operating Bureaus and ten Staff Offices. The Bureaus' responsibilities include: processing applications for licenses and other filings; analyzing complaints; conducting investigations; developing and implementing regulatory programs; and taking part in hearings. http://www.fcc.gov
FEC	Forward Error Correction	A technique of error detection and correction in which a transmitting host computer includes a number of redundant bits in the payload (data field) of a block or frame of data. The receiving device uses the extra bits to detect, isolate and correct any errors created in transmission.
FGS	Fine-Granular Scalability	
FMC	Fixed Mobile Convergence	Convergence between the mobile and fixed line networks giving telecommunications operators the possibility to provide services to users irrespective of their location, access technology, and terminal.
FR	Full Reference	

FTP	File Transfer Protocol	A communication protocol mainly used on Internet to transfer files and make repositories dedicated to file exchange (instead of displaying it directly to the screen). Specified by IETF in RFC 959. http://www.ietf.org
GERAN	GPRS/EDGE Radio Access Network	The Radio Access part of GSM/EDGE. More specifically: RF layer, Layer 1, 2 and 3, internal (Abis, Ater) and external (A, Gb) interfaces, conformance test specifications for all aspects of GERAN base stations and terminals and GERAN specific O&M specifications for the nodes in the GERAN. Specified by 3GPP. http://www.3gpp.org
GGSN	Gateway GPRS Support Node	Interface between the GPRS wireless data network and other networks such as the Internet or private networks. It supports the edge routing function of the GPRS network. To external packet data networks the GGSN performs the task of an IP router. Firewall and filtering functionality, to protect the integrity of the GPRS core network, are also associated with the GGSN along with a billing function.
GIPS	Global IP Sound	A Swedish/American company that develops embedded voice processing solutions for real-time communications on packet networks. The product portfolio includes voice codecs that are robust against packet loss. http://www.globalipsound.com
GoP	Groups of Pictures	
GPRS	General Packet Radio Service	An enhancement to the GSM mobile communication system that supports data packets. GPRS enables continuous flows of IP data packets over the system for such applications as web browsing and file transfer. Supports up to 160 kb/s gross transfer rate. Practical rates are from 12 to 48 kb/s. http://www.etsi.org , http://www.3gpp.org
GRX	GPRS Roaming eXchange	
GSM	Global System for Mobile communications	A digital cellular phone technology system that is the predominant system in Europe, but is also used around the world. Development started in 1982 by CEPT and was transferred to the new organisation ETSI in 1988. Originally, the acronym was the group in charge, "Group Special Mobile" but later the group changed name to SMG. GSM was first deployed in seven countries in Europe in 1992. It operates in the 900 MHz and 1.8 GHz band in Europe and 1.9 GHz band in North America. GSM defines the entire cellular system, from the air interface to the network nodes and protocols. As of January 2005, there were more than 1.2 billion GSM users in more than 200 countries worldwide. The ubiquity of the GSM standard makes international roaming very common between mobile phone operators which enables phone users to access their services in many other parts of the world as well as their own country. GSM differs significantly from its predecessors in that both signalling and speech channels are digital, which means that it is seen as a second generation (2G) mobile phone system. This fact has also meant that data communication was built into the system from very early on. GSM is an open standard which is currently developed by the 3GPP. http://www.gsmworld.com/ , http://www.etsi.org , http://www.3gpp.org
GSM BSS	GSM Base Station Subsystem	
GSMA	GSM Association	World's leading wireless industry representative body, consisting of more than 660 second- and third-generation wireless network operators and key manufacturers and suppliers to the wireless industry. http://www.gsmworld.com/
GTP	GPRS Tunnelling Protocol	An IP based protocol used within GSM and UMTS networks. The GTP protocol is layered on top of UDP. There are in fact three separate protocols, GTP-C, GTP-U and GTP'. GTP-C is used within the GPRS core network for signalling between GPRS Support Nodes (GGSNs and SGSNs). This allows the SGSN to activate a session on the user's behalf (PDP context activation), to deactivate the same session, to adjust quality of service parameters or to update a session for a subscriber who has just arrived from another SGSN. GTP-U is used for carrying user data within the GPRS core network and between the Radio Access Network and the core network. The user data transported can be packets in any of IPv4, IPv6 or PPP formats. GTP' (GTP prime) uses the same message structure as GTP-C and GTP-U, but it is an almost completely separate protocol. It can be used for carrying charging data from the "Charging Data Function" of the GSM or UMTS network to the "Charging Gateway Function".

GW	Gateway	A network element equipped for interfacing with another network that uses different protocols (e.g. Between an IP network and PSTN). Also called Interworking unit/function – IWU/IWF.
HAWAII	Handoff-Aware Wireless Access Internet Infrastructure	
HC	Hybrid Controller	
HCCA	HCF Controlled Channel Access	
HCF	Hybrid Coordination Function	Hybrid of Distributed Coordination Function (DCF) and Point Coordination Function (PCF) used in Wi-Fi Wireless LANs.
HD-SDI	High-Definition Serial Digital Interface	The Serial Digital Interface (SDI) standard is defined by the Society of Motion Picture and Television Engineers (SMPTE), widely used in the broadcasting and video production industry today. SDI standard describes how to carry uncompressed serial, digitized video data between equipment in production facilities over video coax cables. There are two variations of SDI standard based on the data rate: standard-definition (SD)-SDI and high-definition (HD)-SDI. The basic electrical specifications of these two variations are the same, but the main difference is that HD-SDI has a higher data rate at 1.485 Gb/s and 1.485/1001 Gb/s while the SD-SDI data rate ranges from 143 Mb/s to 540 Mb/s, with 270 Mb/s being the most popular rate. http://www.smppte.org/
HDTV	High Definition Television	Broadcast of television signals with a higher resolution than traditional formats (NTSC, SECAM, PAL) allow. Except for an early analog format in Japan, HDTV is broadcast digitally, and therefore its introduction sometimes coincides with the introduction of digital television (DTV). An HDTV-compatible TV usually uses a 16:9 aspect ratio. The high resolution images (1920 pixels × 1080 lines or 1280 pixels × 720 lines) allow much more detail to be shown compared to analog television or regular DVDs. MPEG-2 is currently used as the compression codec. Like NTSC and PAL, 1920 × 1080 broadcasts generally use interlacing to reduce bandwidth demands. Alternating scan lines are broadcast 50 or 60 times a second, similar to PAL's 50 Hz and NTSC's 60 Hz interlacing. This format is entitled 1080i, or 1080i60. In areas traditionally using PAL 50 Hz 1080i50 is also used. Progressive scan formats are also used with frame rates up to 60 per second. The 1280 × 720 format is in practice always progressive scan (with the entire frame refreshed each time) and is thus termed 720p.
HLR/AUC	Home Location Register / Authentication Centre	The Home Location Register or HLR is a central database that contains details of each mobile phone subscriber that is authorized to use the GSM core network. More precisely, the HLR stores details of every SIM card issued by the mobile phone operator. Each SIM has a unique identifier called an IMSI which is one of the primary keys to each HLR record. The next important items of data associated with the SIM are the telephone numbers used to make and receive calls to the mobile phone, known as MSISDNs. The main MSISDN is the number used for making and receiving voice calls and SMS, but it is possible for a SIM to have other secondary MSISDNs associated with it for fax and data calls. Each MSISDN is also a primary key to the HLR record. http://www.etsi.org
HSS	Home Subscriber Service (previous HLR)	
HTTP	Hyper Text Transport Protocol	An application-level protocol for distributed, collaborative, hypermedia information systems. Used to request and transmit files, especially webpages and webpage components, over the Internet or other computer networks. http://www.w3c.org
IANA	Internet Assigned Numbers Authority	IANA (Internet Assigned Numbers Authority) is the organization under the Internet Architecture Board (IAB) of the Internet Society that, under a contract from the US government, has overseen the allocation of Internet Protocol addresses to Internet service providers (ISPs). IANA has also had responsibility for the registry for any "unique parameters and protocol values" for Internet operation. These include port numbers, character sets, and MIME media access types. Partly because the Internet is now a global network, the US government has withdrawn its oversight of the Internet, previously contracted out to IANA, and lent its support to a newly-formed organization with global, non-government representation, the Internet Corporation for Assigned Names and Numbers (ICANN). ICANN has now assumed responsibility for the tasks formerly performed by IANA. http://www.iana.org/
IAPP	Inter-Access Point Protocol	

IBCF	Interconnect Border Control Function	
ICS	Implementation Conformance Statement	
I-CSCF	Interrogating CSCF	A node of the IP Multimedia Subsystem (IMS) specified by 3GPP. A SIP proxy located at the edge of an administrative domain. Its IP address is published in the DNS of the domain (using NAPTR and SRV type of DNS records), so that remote servers (e.g. a P-CSCF in a visited domain, or an S-CSCF in a foreign domain) can find it, and use it as an entry point for all SIP packets to this domain. The I-CSCF queries the HSS using the DIAMETER Cx and Dx interfaces to retrieve the user location, and then route the SIP request to its assigned S-CSCF. http://www.3gpp.org , http://www.ietf.org
ID	Identity	
IDS	Intrusion Detection System	A software/hardware tool used to detect unauthorised access to a computer system or network. This may take the form of attacks by skilled malicious hackers, or Script kiddies using automated tools. An IDS is required to detect all types of malicious network traffic and computer usage. This includes network attacks against vulnerable services, data driven attacks on applications, host based attacks such as privilege escalation, unauthorised logins and access to sensitive files, and malware (viruses, Trojan horses, and worms).
IEC	International Electrotechnical Commission	An organization that sets international electrical and electronics standards founded in 1906. It is made up of national committees from over 40 countries. Headquarters in Geneva, Switzerland. http://www.iec.ch
IEEE	The Institute of Electrical and Electronics Engineers	USA based organisation open to engineers and researchers in the fields of electricity, electronics, computer science and telecommunications. Established in 1884. The aim is to promote research through journals and conferences and to produce standards in telecommunications and computer science. IEEE has produced more than 900 active standards and has more than 700 standards under development. Divided into different branches, or 'Societies'. Has daughter organisations, or 'chapters' in more than 175 countries worldwide. Headquarters in Piscataway, New Jersey, USA. http://www.ieee.org
IETF	Internet Engineering Task Force	A large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet. It is open to any interested individual. The technical work of the IETF is done in its working groups, which are organized by topic into several areas (e.g. routing, transport, security, etc.). Much of the work is handled via mailing lists. The IETF holds meetings three times per year. The IETF working groups are grouped into areas, and managed by Area Directors (AD). The ADs are members of the Internet Engineering Steering Group (IESG). Providing architectural oversight is the Internet Architecture Board (IAB). The IAB also adjudicates appeals when someone complains that the IESG has failed. The IAB and IESG are chartered by the Internet Society (ISOC) for these purposes. The General Area Director also serves as the chair of the IESG and of the IETF, and is an ex-officio member of the IAB. The Internet Assigned Numbers Authority (IANA) is the central coordinator for the assignment of unique parameter values for Internet protocols. The IANA is chartered by the Internet Society (ISOC) to act as the clearinghouse to assign and coordinate the use of numerous Internet protocol parameters. www.ietf.org/rfc/rfc3935.txt
iLBC	Internet Low Bitrate Coder	A GlobalIPSound speech coding algorithm standardised in IETF RFC 3951. http://www.globalipsound.com , http://www.ietf.org
IM	Instant Messaging	An instant messaging service is reached by the use of an instant messenger client. Instant messaging differs from e-mail in that conversations happen in real-time. Also, most services convey an "online status" between users, such as if a contact is actively using the computer. Generally, both parties in the conversation see each line of text right after it is typed (line-by-line), thus making it more like a telephone conversation than exchanging letters. Instant messaging applications may also include the ability to post an away message, the equivalent of the message on a telephone answering machine. Popular instant messaging services on the public Internet include Jabber, AOL Instant Messenger, Yahoo! Messenger, .NET Messenger Service and ICQ. These services owe many ideas to an older (and still popular) online chat medium known as Internet Relay Chat (IRC).
IM	IP Multimedia	

IM-MG	IP Multimedia – Media Gateway	
IMS	IP Multimedia Subsystem	<p>The IP Multimedia Subsystem (IMS) is a standardised Next Generation Networking (NGN) architecture for telecom operators that want to provide mobile and fixed multimedia services. It uses a Voice-over-IP (VoIP) implementation based on a 3GPP standardised implementation of SIP, and runs over the standard Internet Protocol (IP). Existing phone systems (both packet-switched and circuit-switched) are supported. IMS was originally defined by an industry forum called 3G.IP (www.3gip.org) formed in 1999. 3G.IP developed the initial IMS architecture, which was brought to 3GPP for industry standardization as part of their standardization work for 3G mobile phone systems in UMTS networks. It first appeared in release 5 (evolution from 2G to 3G networks), when SIP-based multimedia was added. Support for the older GSM and GPRS networks was also provided. “Early IMS” was defined to allow for IMS implementations that do not yet support all “Full IMS” requirements. 3GPP2 (a different organisation) based their CDMA2000 Multimedia Domain (MMD) on 3GPP IMS, adding support for CDMA2000.</p> <p>http://www.3gpp.org, http://www.ietf.org</p>
IM-SSF	IP Multimedia – Service Switching Function	
IN	Intelligent Networks	<p>The enhanced public switched telephone network architecture for the 1990s developed by ITU. It was created to provide a variety of advanced telephony services such as 800 number translation, local number portability (LNP), call forwarding, call screening and wireless integration. The IN uses the SS7 signalling protocol in which voice calls (or modem data) travel through circuit-switched voice switches, while control signals travel over an SS7 packet-switched network.</p> <p>http://www.itu.int</p>
IP	Internet Protocol	<p>A protocol for communication between computers, used as a standard for transmitting data over networks and as the basis for standard Internet protocols.</p> <p>http://www.ietf.org</p>
IP-CAN	IP Connectivity Access Network	
IPPM	IP Performance Metrics	
IPSec	IP Security	<p>Security protocol for the internet.</p> <p>http://www.ietf.org</p>
IPv4	Internet Protocol v4	<p>IPv4 is version 4 of the Internet Protocol (IP) and it is the first version of the Internet Protocol to be widely deployed. IPv4 is the dominant network layer protocol on the internet. It is described in IETF RFC 791 (September 1981) which obsoleted RFC 760 (January 1980). IPv4 is a data-oriented protocol to be used on a packet switched internetwork (e.g. Ethernet). It is a best effort protocol in that it does not guarantee delivery. It does not make any guarantees on the correctness of the data; it may result in duplicated packets and/or packets out-of-order. All of these things are addressed by an upper layer protocol (e.g. TCP, UDP). See also IP and Ipv6.</p> <p>http://www.ietf.org</p>
IPv6	Internet Protocol v6	<p>Ipv6 is version 6 of the Internet Protocol (IP). A network layer standard used by electronic devices to exchange data across a packet-switched internetwork. It follows IPv4 as the second version of the IP to be formally adopted for general use. IPv6 is intended to provide more addresses for networked devices, allowing, for example, each cell phone and mobile electronic device to have its own address. IPv4 supports 4.3 billion addresses, which is inadequate to give one (or more if they possess more than one device) to every living person. IPv6 supports 3.4×10^{38} addresses, or 5×10^{28} (50 octillion) for each of the roughly 6.5 billion people alive today. Invented by Steve Deering and Craig Mudge at Xerox PARC, IPv6 was adopted by the Internet Engineering Task Force in 1994, when it was called “IP Next Generation” (IPng). As of December 2005, IPv6 accounts for a tiny percentage of the live addresses in the publicly-accessible Internet, which is still dominated by IPv4. The adoption of IPv6 has been slowed by the introduction of network address translation (NAT), which partially alleviates address exhaustion.</p> <p>http://www.ietf.org</p>
IRAP	International Roaming Access Protocol	<p>An emerging protocol that enables operation across different types of wireless and wired networks. It also provides the seamless connectivity across the heterogeneous network.</p>
ISC	IP Multimedia Subsystem Service Control	

ISDN	Integrated Services Digital Network	A digital telecommunications network that provides end-to-end digital connectivity to support a wide range of services, including voice and non-voice services, to which users have access by a limited set of standard multi-purpose user-network interfaces. The user is offered one or more 64 kb/s channels. http://www.itu.int
ISDN BRI	ISDN Basic Rate Interface	A user-network access arrangement that corresponds to the interface structure composed of two B-channels and one D-channel. The bit rate of the D-channel for this type of access is 16 kbit/s. http://www.itu.int
ISDN PRI	ISDN Primary Rate Interface	A user-network access arrangement that corresponds to the interface structure composed of 30 B-channels and one D-channel. The bit rate of the D-channel for this type of access is 64 kbit/s. http://www.itu.int
ISO	International Standardisation Organisation	The International Organization for Standardization (ISO) is a worldwide federation of national standards bodies from more than 140 countries, one from each country. ISO is a non-governmental organization established in 1947. The mission of ISO is to promote the development of standardization and related activities in the world with a view to facilitating the international exchange of goods and services, and to developing cooperation in the spheres of intellectual, scientific, technological and economic activity. http://www.iso.org
ISP	Internet Service Provider	A vendor who provides access for customers to the Internet and the World Wide Web. The ISP also typically provides a core group of internet utilities and services like e-mail and news group readers.
ISUP	Integrated Services Digital Network – User Part	This encompasses the signalling functions in SS No. 7 required to provide switched services and user facilities for voice and non-voice applications in an ISDN (see Recommendation Q.761). http://www.itu.int
ITU-R	International Tele-communication Union – Radio communication Sector	http://www.itu.int/ITU-R
ITU-T	International Tele-communication Union – Standardization Sector	http://www.itu.int/ITU-T/
IWF	Interworking Function	A function connecting two networks of differing signalling technology or administrative policies. See also Gateway (GW) or Interworking Unit (IWU). http://www.etsi.org/Teddi
IWU	Interworking Unit	A network element supporting IWFs. See also Interworking Function (IWF) or Gateway (GW). http://www.etsi.org/Teddi
LAN	Local Area Network	A network shared by communicating devices, usually on a small geographical area. A system that links together electronic office equipment, such as computers and word processors, and forms a network within an office or building.
LBR	Low Bit-rate Redundancy	
LDAP	Lightweight Directory Access Protocol	A networking protocol for querying and modifying directory services running over TCP/IP. Its current version is LDAPv3, as defined in RFC 3377. http://www.ietf.org
LDAS	Low Delay Audio Streamer	
LD-CELP	Low Delay Code-Excited Linear Prediction	A speech coding technology. It is used by ITU-T in the G.728 standard for speech coding operating at 16 kb/s. Delay of the codec is only 5 samples (0.625 ms). The linear prediction is calculated backwards with a 50th order LPC filter. The excitation is generated with gain scaled VQ. The standard was finished in 1992 in the form of algorithm exact floating point code. In 1994 a bit exact fixed point codec was released. G.728 passes low bit rate modem signals up to 2400 bit/s. Also network signalling goes through. The complexity of the codec is 30 MIPS. 2 kBytes of RAM is needed for codebooks.
LEAF	Linux Embedded Appliance Firewall	

LPC	Linear Predictive Coding	A speech coding algorithm. It is a tool used mostly in audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form, using the information of a linear predictive model. It is a powerful speech analysis technique, and one of the most useful methods for encoding good quality speech at a low bit rate and provides extremely accurate estimates of speech parameters.
MAC	Medium Access Control	The lower of the two sub layers of the Data Link Layer. In general terms, MAC handles access to a shared medium, and can be found within many different technologies. For example, MAC methodologies are employed within Ethernet, GPRS, and UMTS.
MAP	Mobile Application Part	A protocol that enables real time communication between nodes in a mobile cellular network. A typical usage of the MAP protocol would be for the transfer of location information from the VLR (Visitor Location Register) to the HLR (Home Location Register).
MBone	Multicast Backbone	
MCU	Multipoint Control Unit	
MD5	Message Digest Algorithm 5	
MECCANO	Multimedia Education and Conferencing Collaboration over ATM Networks and Others	A European Union project from 1998 to 2000 under the Telematics Application Programme. The objective of the project was to provide all the technology components, other than the data network itself, to support collaborative research and technical development through the deployment of enhanced tools for multimedia collaboration in Europe. It succeeded the MERCI project. http://www-mice.cs.ucl.ac.uk/multimedia/projects/meccano/
MEGACO	Media Gateway Control	An IETF WG that jointly with ITU-T SG 16 has developed a Gateway Control Protocol. The IETF version is RFC 3525, while the ITU-T version is specified in ITU-T Recommendation H. 248.1. ITU-T has continued the work on the protocol, there are a number of packages that define additional functionality. http://www.ietf.org
MERCI	Multimedia European Research Integration	A European Union research project running from 1995 to 1997 under the Telematics Application Programme. The objective of the project is to provide all the technology components, other than the data network itself, to allow proper deployment of the tools for European multimedia collaboration in Europe. It was succeeded by the MECCANO project. http://www-mice.cs.ucl.ac.uk/multimedia/projects/merci/
MF	Multi Field	
MG	Media Gateway	A Media Gateway acts as a translation unit between disparate telecommunications networks such as PSTN; Next Generation Networks; 2G, 2.5G and 3G radio access networks or PBX. Media Gateways enable multimedia communications across Next Generation Networks over multiple transport protocols such as ATM and IP.
MGC	Media Gateway Controller	Controls the Media Gateways. http://webapp.etsi.org/Teddi/
MGCF	Media Gateway Control Function	The functions of a Media Gateway Controller
MGW	Media Gateway	Converts media provided in one type of network to the format required in another type of network. http://webapp.etsi.org/Teddi/
MH	Mobile Host	
MIB	Management Information Base	
MICE	Multi-media Integrated Conferencing for Europe	
MIME	Multipurpose Internet Mail Extensions	Multipurpose Internet Mail Extensions (MIME) is an Internet Standard for the format of e-mail. Virtually all Internet e-mail is transmitted via SMTP in MIME format. Internet e-mail is so closely associated with the SMTP and MIME standards that it is sometimes called SMTP/MIME e-mail.
MLT	Modulated Lapped Transform (speech coding)	A speech coding algorithm.
MMoIP	Multimedia over IP	

MMS	Multimedia Message Service	MMS – sometimes called Multimedia Messaging System – is a communications technology developed by 3GPP (Third Generation Partnership Project) that allows users to exchange multimedia communications between capable mobile phones and other devices. An extension to the Short Message Service (SMS) protocol, MMS defines a way to send and receive, almost instantaneously, wireless messages that include images, audio, and video clips in addition to text. http://www.3gpp.org
MOS	Mean Opinion Score	A numerical indication of the perceived quality of received human speech over the connection. The MOS is expressed as a single number in the range 1 to 5, where 1 is lowest perceived quality, and 5 is the highest perceived quality. MOS tests are specified by ITU-T Recommendation P.800. The MOS is generated by averaging the results of a set of standard, subjective tests where a number of listeners rate the perceived audio quality of test sentences read aloud by both male and female speakers over the communications medium being tested. http://www.itu.int
MP3	MPEG-1/2 Audio Layer-3	A standard technology and format for compressing a sound sequence into a very small file (about one-twelfth the size of the original file) while preserving the original level of sound quality when it is played. MP3 files (identified with the file name suffix of “.mp3”) are available for downloading from a number of websites. MP3 files are usually download-and-play files rather than streaming sound files that you link-and-listen-to with RealPlayer and similar products. (However, streaming MP3 is possible.)
MPE	Multi-Pulse Excited	A speech coding algorithm.
MPEG	Moving Pictures Experts Group	MPEG is a working group of ISO/IEC (ISO/IEC JTC/SC29 WG 11) in charge of the development of standards for coded representation of digital audio and video. The working group has developed three sets of standards defining coding and transmission of audio and video; MPEG-1, MPEG-2 and MPEG-4. Another standard developed is MPEG-7, the standard for description and search of audio and visual content. MPEG-21 is a standard defining Multimedia Framework. http://www.chiariglione.org/mpeg/
MPLS	Multi Protocol Label Switching	An IETF standard intended for Internet application. MPLS has been developed to speed up the transmission of IP based communications over ATM networks. The system works by adding a much smaller “label” to a stream of IP datagrams allowing “MPLS” enabled ATM switches to examine and switch the packet much faster. It is specified in IETF RFC 2702. www.ietf.org
MP-MLQ	Multipulse Maximum Likelihood Quantization	A speech coding algorithm.
MRF	Multimedia Resource Function	
MRFC	Multimedia Resources Function Controller	
MRFP	Multimedia Resources Function Processor	
MSC	Mobile services Switching Centre	The Mobile services Switching Centre or MSC is a sophisticated telephone exchange which provides circuit-switched calling, mobility management and GSM services to the mobile phones roaming within the area that it serves. This means voice, data and fax services, as well as SMS and call divert. A Gateway MSC is the MSC that determines which visited MSC the subscriber who is being called is currently located in. It also interfaces with the Public Switched Telephone Network. All mobile to mobile calls and PSTN to mobile calls are routed through a GMSC. The term is only valid in the context of one call since any MSC may provide both the gateway function and the Visited MSC function; however, some manufacturers design dedicated high capacity MSCs which do not have any BSCs connected to them. These MSCs will then be the Gateway MSC for many of the calls they handle. http://www.etsi.org
MSDU	MAC Service Data Unit	Information that is delivered as a unit between MAC service access points (SAPs) of a IEEE 802-11 WLAN. http://www.ieee802.org
MSN	Microsoft Network	http://www.microsoft.com
MSN	Multiple Subscriber Number	An ISDN supplementary service. http://www.itu.int

MT	Mobile Terminal	
MTP	SS7 Message Transfer Part	The Message Transfer Part (MTP) is part of the Signalling System 7 (SS7) used for communication in Public Switched Telephone Networks (PSTN). MTP is responsible for the correct and reliable end to end data transport of SS7 messages between communication partners. In the OSI model, MTP Level 2 corresponds to OSI Layer 2 (data link layer) and MTP Level 3 to the OSI Layer 3 (network layer). MTP is formally defined in ITU-T recommendations Q.701-Q.705. http://www.itu.int
NAL	Network Adaptation Layer	
NALU	NAL Unit	
NAPT	Network Address and Protocol Translator	
NAPT-PT	Network and Port Translation – Protocol Translation	
NAPTR	Naming Authority Pointer	
NASA	National Aeronautic and Space Administration	A US government agency established in 1958, partially in response to the Soviet Union's launch of the first artificial satellite. Its mission is to perform research and exploration of space and space technology. http://www.nasa.gov
NASS	Network Attachment Subsystem	
NAT	Network Address Translation	In computer networking, the process of network address translation (NAT, also known as network masquerading or IP-masquerading) involves re-writing the source and/or destination addresses of IP packets as they pass through a router or firewall. Most systems using NAT do so in order to enable multiple hosts on a private network to access the Internet using a single public IP address. According to specifications, routers should not act in this way, but many network administrators find NAT a convenient technique and use it widely. Nonetheless, NAT can introduce complications in communication between hosts. NAT first became popular as a way to deal with the IPv4 address shortage and to avoid the difficulty of reserving IP addresses. Use of NAT has proven particularly popular in countries other than the United States, which (for historical reasons) have fewer address-blocks allocated per capita. It has become a standard feature in routers for home and small-office Internet connections, where the price of extra IP addresses would often outweigh the benefits. In a typical configuration, a local network uses one of the designated "private" IP address subnets (such as 192.168.x.x or 10.x.x.x), and a router on that network has a private address (such as 192.168.0.1) in that address space. The router is also connected to the Internet with a single "public" address (known as "overloaded" NAT) or multiple "public" addresses assigned by an ISP. As traffic passes from the local network to the Internet, the source address on the packets are translated on the fly from the private addresses to the public address(es). The router tracks basic data about each active connection (particularly the destination address and port). When a reply returns to the router, it uses the connection tracking data it stored during the outbound phase to determine where on the internal network to forward the reply; the TCP or UDP client port numbers are used to demultiplex the packets in the case of overloaded NAT, or IP address and port number when multiple public addresses are available, on packet return. To a system on the Internet, the router itself appears to be the source/destination for this traffic. The use of NAT creates problems for applications where the source and destination addresses and port numbers are used in the protocols, such as voice over IP.
NAT-T	NAT Transversal	
NAV	Network Allocation Vector	
NC	Network Control	
netlmm	Network-based Localized Mobility Management	An IETF Working Group. The task is to design a protocol solution for network-based localized mobility management. http://www.ietf.org

NGN	Next Generation Network	A network concept that aims at providing a framework to encompass the large variety of existing and emerging protocols and services, facilitate a further evolution of these, decouple the evolution from the underlying network infrastructure and facilitate the interfacing of a plethora of available media. The rationale behind NGN lies founded in paradigm shifts that have been taking place within the technological solutions and the business models in the telecom industry as a whole. The concept is based on IP-technology and is being specified by ITU-T. http://www.itu.int
NPT	Norwegian Post and Telecommunications Authority	Norwegian Post and Telecommunications Authority (NPT) is an autonomous administrative agency under the Norwegian Ministry of Transport and Communications, with monitoring and regulatory responsibilities for the postal and telecommunications markets in Norway. The NPT is self-financed, primarily through fees and charges. http://www.npt.no
NR	No-Reference	
NS	Notification Server	
NTNU	Norwegian University of Science and Technology	Part of the University in Trondheim (UNiT), Norway. It was established in 1996 as a further development of UNiT as a result of merger between The Norwegian Institute of Technology (NTH), The College of Arts and Sciences (AVH) and the Museum of Natural History and Archaeology (VM). It offers Professional degrees, university studies, interdisciplinary study programmes and Masters degrees in English. It has seven faculties and 53 departments. It has 20,000 students, half of these studying technology or the natural sciences. www.ntnu.no
OAM	Operations, Administration and Maintenance	
OAN	Open Access Network	Network model which refers to a horizontally layered network architecture and business model that separates physical access to the network from service provisioning. The same OAN will be used by a number of different providers that share the investments and maintenance cost. The OAN concept is especially appropriate for deploying metropolitan Wi-Fi Access Networks.
OFDM	Orthogonal Frequency Division Multiplexing	A spread spectrum technique that distributes the data over a large number of carriers spaced apart at precise frequencies. This spacing provides the "orthogonality" in this technique, which prevents the demodulators from seeing frequencies other than their own. The benefits of OFDM are high spectral efficiency, resiliency to RF interference, and lower multipath distortion. This is useful because in a typical terrestrial wireless scenario there are multipath channels (i.e. the transmitted signal arrives at the receiver using various paths of varying length). Since multiple versions of the signal interfere with each other (inter symbol interference (ISI)) it becomes very hard to extract the original information. OFDM is sometimes called multi-carrier or discrete multi-tone modulation. It is the modulation technique used for digital TV in Europe, Japan and Australia. It is used in DAB, ADSL and WLAN 802.11a and g and WMAN 802.16 standards.
OSA	Open Service Access	
OSA SCS	OSA Service Capability Server	
P2P	Peer To Peer	A computer network that does not rely on dedicated servers for communication but instead mostly uses direct connections between clients (peers). A pure peer-to-peer network does not have the notion of clients or servers, but only equal peer nodes that simultaneously function as both "clients" and "servers" to the other nodes in the network.
PABX	Private Automatic Branch Exchange	Also called PBX or Private Business eXchange. A telephone exchange that is owned by a private business, as opposed to one owned by a common carrier or by a telephone company.
PATS	Publicly Available Telephone Services	
PC	Personal Computer	Usually a microcomputer whose price, size, and capabilities make it suitable for personal usage. Personal computers are normally operated by one user at a time to perform such general purpose tasks as word processing, internet browsing, e-mail and other digital messaging, multimedia playback, video game play, computer programming, etc. Unlike many special purpose and high performance computers, it is assumed that a typical personal computer will run software not written by its primary users.
PC	Point Coordinator	A device that carries out the Point Coordination Function in an IEEE 802.11 WLAN.

PCF	Point Coordination Function	Point Coordinated Function is a Medium Access Control (MAC) technique used in wireless networks which relies on a central node, often an Access Point (AP), to communicate with a node listening, to see if the radio resource is free.
PCF	Policy Control Function	
PCM	Pulse Code Modulation	The speech coding algorithm used in most circuit-switched fixed networks. PCM is a wave form coding method which is neutral to the actual content of the signal. See also A/D and D/A.
P-CSCF	Proxy CSCF	A SIP proxy that is the first point of contact for an IP Multimedia Subsystem (IMS) terminal. It can be located either in the visited network (in full IMS networks) or in the home network (when the visited network is not IMS compliant yet). http://www.3gpp.org , http://www.ietf.org
PDA	Personal Digital Assistant	Handheld device that combines computing, telephone/fax, Internet and networking features. A typical PDA can function as a cellular phone, fax sender, Web browser and personal organizer.
PDF	Policy Decision Function	
PDP	Packet Data Protocol	A packet transfer protocol used in wireless GPRS networks. http://www.3gpp.org
PDP	Policy Decision Point	
PEAQ	Perceptual Evaluation of Audio Quality	
PEP	Policy Enforcement Point	
PES	PSTN/ISDN Emulation Subsystem	
PESQ	Perceptual Evaluation of Speech Quality	
PHB	Per-Hop-Behaviour	The externally observable forwarding behaviour applied at a DiffServ-compliant node to a DiffServ behaviour aggregate. IETF RFC 3564. http://www.ietf.org
PHY	Physical layer device	The Ethernet PHY at Layer 1 of the OSI model defines the electrical and optical signalling, line states, clocking guidelines, data encoding, and circuitry needed for data transmission and reception. Contained within the PHY are several sub-layers that perform these functions including the physical coding sub-layer (PCS) and the optical transceiver or physical media dependent (PMD) sub-layer for fibre media. The Ethernet PHY connects the media to the MAC (Layer 2).
PIB	Policy Information Base	
PIFS	PCF Interframe Space	
PIN	Personal Identity Number	
PLC	Packet Loss Concealment	
PLMN	Public Land Mobile Network	Common notation in the 80s of a land mobile network of any category that was used to offer public services.
POTS	Plain Old Telephone Service	A very general term used to describe an ordinary voice telephone service. See also PSTN.
PPPoE	Point to Point Protocol over Ethernet	A network protocol for encapsulating PPP frames in Ethernet frames. It is used mainly with DSL services. It offers standard PPP features such as authentication, encryption, and compression. PPPoE is a tunnel protocol which allows one to layer IP over a connection between two Ethernet ports, but with the software features of a PPP link, so it is used to virtually "dial" to another Ethernet machine and make a point to point connection with it, which is then used to transport IP packets, based on the features of PPP. It is specified by IETF RFC 2516. http://www.ietf.org
PRI	Primary Rate Interface (ISDN)	See ISDN PRI.

PS	Packet Switched	Communication switching method in which packets (units of information carriage) are individually routed between nodes over data links which might be shared by many other nodes. Packet switching is used to optimize the use of the bandwidth available in a network, to minimize the transmission latency (i.e. the time it takes for data to pass across the network), and to increase robustness of communication. The concept of packet switching was developed by Paul Baran in the early 1960s, and independently a few years later by Donald Davies, as described below. Leonard Kleinrock conducted early research and published a book in the related field of digital message switching (without the packets) in 1961, and also later played a leading role in building and management of the world's first packet switched network, the ARPANET.
PSTN	Public Service Telephone Network	Common notation for the conventional analog telephone network.
Q2S	Center for Quantifiable quality of Service in Communication Systems	A Norwegian Centre of Excellence at the Norwegian University of Science and Technology (NTNU) in Trondheim. The Centre will study principles, derive mechanisms, methods and technical solutions and assess their properties and performances by means of experiments and models. Performances relate to perceived quality of streamed speech/music and video, delays and throughput of elastic traffic, reliability and availability of services, and information security with encryption and user authentication. http://www.q2s.ntnu.no/
QAP	QoS Access Point	
QoE	Quality of Experience	User's perceived experience of what is being presented by a communication service or application user interface. http://webapp.etsi.org/Teddi/
QoS	Quality of Service	The "degree of conformance of the service delivered to a user by a provider, with an agreement between them". The agreement is related to the provision/delivery of this service. Defined by EURESCOM project P806 in 1999 and adopted by ITU-T in Recommendation E.860 [E.860]. http://www.itu.int , http://www.eurescom.de
QSTA	QoS Station	
RAC	Resource and Admission Control	
RACS	Resource and Admission Control Subsystem	
RAN	Radio Access Network	A part of a mobile telecommunication system. It implements a radio access technology. Conceptually, it sits between the mobile phone and the core network (CN). Examples are GERAN (GSM RAN), GERAN (GSM/EDGE RAN) and UTRAN (UMTS RAN).
RAP	Resource Allocation Protocol	
RAS	Registration, Admission, Status	
RC4	Rivest Cipher 4	One of several ciphers proprietary to the RSA Data Security Inc. Also called ARCFOUR, it is the most widely-used software stream cipher and is used in popular protocols such as Secure Sockets Layer (SSL) (to protect Internet traffic) and WEP (to secure wireless networks). RC4 falls short of the high standards of security set by cryptographers, and some ways of using RC4 lead to very insecure cryptosystems (including WEP). RC4 was designed by Ron Rivest of RSA Security in 1987; while it is officially termed "Rivest Cipher 4", the RC acronym is alternatively understood to stand for "Ron's Code". RC4 was initially a trade secret, but in September 1994 a description of it was anonymously posted to the <i>Cypherpunks</i> mailing list. It was soon posted on the <i>sci.crypt</i> newsgroup, and from there to many sites on the Internet. Because the algorithm is known, it is no longer a trade secret.
RCF	Registration Confirmation	
RFC	Request For Comment	A Request for Comments (RFC) is a formal document from the Internet Engineering Task Force (IETF) that is the result of committee drafting and subsequent review by interested parties. Some RFCs are informational in nature. Of those that are intended to become Internet standards, the final version of the RFC becomes the standard and no further comments or changes are permitted. Change can occur, however, through subsequent RFCs that supersede or elaborate on all or parts of previous RFCs http://www.whatis.com
RGW	Residential Gateway	

R-MGF	Residential Media Gateway Function	
Rn	Release n	Some standards committees (e.g. 3GPP and ETSI TISPAN) organise their standards in Releases. R7 (e.g.) is Release 7, the most recent 3GPP Release.
RRJ	Registration Reject	
RRQ	Registration Request	
RSVP	Resource Reservation Protocol	
RTCP	Real-Time Control Protocol	The Real Time Transport Control Protocol is defined in the same RFC as RTP. The header is similar to the RTP header. It defines the periodic transmission of control packets to all participants in the session, using the same distribution mechanism as the data packets. It is specified that the RTcP protocol should use an odd port number that is one higher than the corresponding RTP port number. http://www.ietf.org
RTFM	Real-time Traffic Flow Measurement	
RTP	Real-Time Transport Protocol	The Internet-standard protocol for the transport of real-time data, including audio and video. RTP is used in virtually all voice-over-IP architectures, for videoconferencing, media-on-demand, and other applications. A thin protocol, it supports content identification, timing reconstruction, and detection of lost packets. The protocol is defined in RFC 3550. RTP should use a dynamically allocated even port number. The actual media coding payload is defined in separate RFCs. RTP combines its data transport with a control protocol (RTCP), which makes it possible to monitor data delivery for large multicast networks, see also RTCP. http://www.ietf.org
RTS	Request To Send	
RTSP	Real-Time Streaming Protocol	
RU	Residential User	
S/MIME	Secure MIME	
SAP	Service Access Point	
SB	Switchboard	
SBC	Session Border Controller	
SBLP	Service Based Local Policy	
SCS	OSA Service Capability Server	
S-CSCF	Serving CSCF	The central node of the signalling plane of the IP Multimedia Subsystem (IMS). It is a SIP server but performs session control as well. It is always located in the home network. The S-CSCF uses DIAMETER Cx and Dx interfaces to the HSS to download and upload user profiles – it has no local storage of the user. http://www.3gpp.org , http://www.ietf.org
SCTP	Stream Control Transmission Protocol	A transport layer protocol defined by the IETF Signaling Transport (SIGTRAN) working group. The protocol is defined in RFC 2960, and an introductory text is provided by RFC 3286. http://www.ietf.org
SDI	Serial Digital Interface	The Serial Digital Interface (SDI) standard is defined by the Society of Motion Picture and Television Engineers (SMPTE), widely used in the broadcasting and video production industry today. SDI standard describes how to carry uncompressed serial, digitized video data between equipment in production facilities over video coax cables. There are two variations of SDI standard based on the data rate: standard-definition (SD)-SDI and high-definition (HD)-SDI. The basic electrical specifications of these two variations are the same, but the main difference is that HD-SDI has higher data rate at 1.485 Gb/s and 1.485/1001 Gb/s, while SD-SDI data rate ranges from 143 Mb/s to 540 Mb/s, with 270 Mb/s being the most popular rate. http://www.smpte.org/
SDO	Standards Developing Organization	
SDP	Session Description Protocol	The Session Description Protocol (SDP) describes multimedia sessions for the purpose of session announcement, session invitation and other forms of multimedia session initiation.

SDU	Service Data Unit	
SG	Signalling Gateway	Provides the signalling mediation function between the IP domain and the SCN domain. http://webapp.etsi.org/Teddi/
SGSN	Serving GPRS support node	The Serving GPRS Support Node is an exchange which performs packet switching functions for mobile stations located in a geographical area designated as the SGSN area. http://webapp.etsi.org/Teddi/
SGW	Signalling Gateway	Provides the signalling mediation function between the IP domain and the SCN domain. http://webapp.etsi.org/Teddi/
SIFS	Short Interframe Space	
SIM	Subscriber Identity Module	A subscriber identity module (SIM) is a logical application running on a UICC smartcard. Although the terms UICC and SIM are often interchanged, UICC refers to the physical card, whereas SIM refers to a single application residing in the UICC that collects GSM user subscription information. The SIM provides secure storing of the key identifying a mobile phone service subscriber but also subscription information, preferences and storage of text messages. The equivalence of a SIM in UMTS is a Universal Subscriber Identity Module (USIM).
SIMPLE	SIP for Instant Messaging and Presence Leveraging Extensions	An IETF Working Group.
SIP	Session Initiation Protocol	An IETF Protocol used to set up voice calls over an IP network. Also the name of the IETF WG developing the protocol.
SIPPING	Session Initiation Proposal Investigation	An IETF Working Group.
SLA	Service Level Agreement	A contract between a provider and a customer that guarantees specific levels of performance and reliability at a certain cost. This contract should also define precisely what could be penalties and back-up solutions in case of problems. SLA is especially important to define when an important part of your system or activity relies on third party providers. SLA is also a very good approach for services provided internally to your organisation where you should also have a customer approach concern {source [RFC3272]}.
SLF	Subscription Locator Function	
SLO	Service Level Objective	
SLS	Service Level Specification	
SMP	Significant Market Power	
SMPTE	Society of Motion Picture and Television Engineers	SMPTE is a professional association for enhancing the profession and contributing to the technology of motion picture and television engineering. The SMPTE establishes standards, practices, and guidelines for the motion picture and television industry, including the audio that goes with the motion images. Founded in 1916 as the Society for Motion Picture Engineers (the T for television was added in 1950), the recent decades have seen computer technology become an important part of the association's concerns and preoccupations. http://www.smpite.org/
SMS	Short Message Service	A means by which short messages can be sent to and from digital cellular phones, pagers and other handheld devices. Alphanumeric messages of up to 160 characters can be supported [Newton03].
SNR	Signal-to-Noise Ratio	The power ratio between the useful signal level (C) and the thermal noise level (N). Often expressed in dB.
SRC	Source	
SS7	Signalling System #7	A set of telephony signalling protocols which are used to set up the vast majority of the world's PSTN telephone calls. http://www.itu.int
STA	Station	Term used to denote users and access points in a Wi-Fi Wireless LAN. http://www.ieee802.org
STF	Special Task Force	ETSI temporary team of specialists assigned for specific purposes.

STM-1	Synchronous Transport Module transmission at 155 Mbit/s	
STM-4	Synchronous Transport Module transmission at 622 Mbit/s	
STUN	Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)	STUN is a protocol that allows entities behind a NAT to first discover the presence of a NAT and the type of NAT, and then to learn the address bindings allocated by the NAT. STUN requires no changes to NATs and works with an arbitrary number of NATs in tandem between the application entity and the public Internet.
TCP	Transport Control Protocol	Transport layer protocol defined for the Internet by Vint Cerf and Bob Kahn in 1974. A reliable octet streaming protocol used by the majority of applications on the Internet, it provides a connection-oriented, full-duplex, point-to-point service between hosts. http://www.ietf.org
TDM	Time Division Multiplex	A type of digital multiplexing in which two or more apparently simultaneous channels are derived from a given frequency spectrum, i.e. bit stream, by interleaving pulses representing bits from different channels.
TE	Terminal Equipment	
TELR	Talker Echo Loss Rating	
THIG	Topology Hiding Internetwork Gateway	
TIA	Telecommunications Industry Association	A US Trade association for the ICT industry, representing the communications sector of the Electronics Industry Alliance (EIA). It was started in 1924 by a small group of suppliers to the independent telephone industry, and later became a committee of the US Independent Telephone Association. In 1979 the group became the US Telecommunications Suppliers Association (USTSA). TIA was formed in 1988 when merging USTSA and the Information and Telecommunications Group of EIA. TIA is accredited by the American National Standards Institute (ANSI) as a major contributor of voluntary industry standards. TIA has more than 70 standards formulating groups. http://www.tiaonline.org
TIPHON	Telecommunications and Internet Protocol Harmonization Over Networks	A former ETSI project on IP telephony standardization. Closed in 2005. The work continues in the Technical Committee (TC) TISPAN. http://portal.etsi.org/tispan
TISPAN	Telecommunication and Internet converged Services and Protocols for Advanced Networking	The ETSI core competence centre for fixed networks and for migration from switched circuit networks to packet-based networks with an architecture that can serve in both. TISPAN is responsible for all aspects of standardisation for present and future converged networks including the NGN (Next Generation Network) and including service aspects, architectural aspects, protocol aspects, QoS studies, security related studies, mobility aspects within fixed networks, using existing and emerging technologies. TISPAN is structured as a single technical committee, with core competencies, under which there are Working Groups and Project Teams. http://www.etsi.org , http://portal.etsi.org/tispan
TKIP	Temporal Key Integrity Protocol	The Temporal Key Integrity Protocol is part of the IEEE 802.11i encryption standard for wireless LANs. TKIP is the next generation of WEP, the Wired Equivalency Protocol, which is used to secure 802.11 wireless LANs. TKIP provides per-packet key mixing, a message integrity check and a re-keying mechanism, thus fixing the flaws of WEP.
TLS	Transport Layer Security	Transport Layer Security (TLS) is a protocol that ensures privacy between communicating applications and their users on the Internet. When a server and client communicate, TLS ensures that no third party may eavesdrop or tamper with any message. TLS is the successor to the Secure Sockets Layer (SSL). http://www.whatis.com
ToE	Target of Evaluation	
ToS	Type of Service	
TR	Technical Report	
TS	Technical Specification	

TSB	Telecommunications Systems Bulletin	A standards document issued by TIA/EIA (US).
TURN	Transversal Using Relay NAT	TURN is a simple protocol that allows for an element behind a Network Address Translation (NAT) or firewall to receive incoming data over TCP or UDP connections. It is most useful for elements behind symmetric NATs or firewalls that wish to be on the receiving end of a connection to a single peer. TURN does not allow for users to run servers on well known ports if they are behind a NAT; it supports the connection of a user behind a NAT to only a single peer. In that regard, its role is to provide the same security functions provided by symmetric NATs and firewalls, but to "turn" the tables so that the element on the inside can be on the receiving end, rather than the sending end, of a connection that is requested by the client
TVRA	Threat Vulnerability and Risk Analysis	
TXOP	Transmission Opportunity	
UA	User Agent	Term used in SIP standards. (UA = UAC + UAS)
UAC	User Agent Client	Term used in SIP standards. A user agent client is a logical entity that creates a new request, and then uses the client transaction state machinery to send it. The role of UAC lasts only for the duration of that transaction. In other words, if a piece of software initiates a request, it acts as a UAC for the duration of that transaction. If it receives a request later, it assumes the role of a user agent server for the processing of that transaction.
UAS	User Agent Server	Term used in SIP standards. A user agent server is a logical entity that generates a response to a SIP request. The response accepts, rejects, or redirects the request. This role lasts only for the duration of that transaction. In other words, if a piece of software responds to a request, it acts as a UAS for the duration of that transaction. If it generates a request later, it assumes the role of a user agent client for the processing of that transaction.
UDP	User Datagram Protocol	UDP is an unreliable protocol used as an alternative to TCP. UDP does not support retransmission of lost packets. It is used for media transport because voice and video transmission is delay sensitive.
UE	User Equipment	
UEP	Unequal Error Protection	
UMTS	Universal Mobile Telecommunication System	The European member of the IMT 2000 family of 3G wireless standards. UMTS supports data rates of 144 kb/s for vehicular traffic, 384 kb/s for pedestrian traffic and up to 2 Mb/s in support of in-building services. The standardisation work began in 1991 by ETSI but was transferred in 1998 to 3GPP as a corporation between Japanese, Chinese, Korean and American organisations. It is based on the use of WCDMA technology and is currently deployed in many European countries. The first European service opened in 2003. In Japan NTT DoCoMo opened its "pre-UMTS" service FOMA (Freedom Of Mobile multimedia Access) in 2000. The system operates in the 2.1 GHz band and is capable of carrying multimedia traffic. http://www.3gpp.org/
UNI	User Network Interface	An interface that is used for the interconnection of customer equipment with a network element of the transport network. http://www.itu.int/sancho
UP	User Priority	
UPnP	Universal Plug and Play	Universal Plug and Play (UPnP) is a standard that uses Internet and Web protocols to enable devices such as PCs, peripherals, intelligent appliances, and wireless devices to be plugged into a network and automatically know about each other. With UPnP, when a user plugs a device into the network, the device will configure itself, acquire a TCP/IP address, and use a discovery protocol based on the Internet's Hypertext Transfer Protocol (HTTP) to announce its presence on the network to other devices. http://www.upnp.org
Upstream		Identifies transmission from the user equipment to the network.
URI	Uniform Resource Identifier	A URI is a compact string of characters for identifying an abstract or physical resource. It is defined by IETF RFC 2396. http://www.ietf.org

URL	Uniform Resource Locator	A subset of Uniform Resource Identifiers (URI) that identify resources via a representation of their primary access mechanism (e.g. their network "location"), rather than identifying the resource by name or by some other attribute(s) of that resource. Originally defined by IETF in RFC 1738, later merged with RFC 1808 to RFC 2396 on URN. http://www.ietf.org
USB	Universal Serial Bus	USB is a plug-and-play interface between a computer and add-on devices (such as audio players, joysticks, keyboards, telephones, scanners, and printers). With USB, a new device can be added to your computer without having to add an adapter card or even having to turn the computer off. The USB peripheral bus standard was developed by Compaq, IBM, DEC, Intel, Microsoft, NEC, and Northern Telecom and the technology is available without charge for all computer and device vendors. http://www.whatis.com
USIM	Universal Subscriber Identity Module	See SIM.
UTRAN	UMTS Radio Access Network	Part of the 3G standard UMTS. The UTRAN consists of a set of Radio Network Subsystems (RNS) connected to the Core Network through the Iu-Interface. An RNS consists of a Radio Network Controller (RNC) and a number of base stations called Node Bs. They provide the radio interface Uu towards the User Equipment (UE). Specified by 3GPP. http://www.3gpp.org
VAD	Voice Activity Detection	
VC1	Video Codec 1	A video codec developed by Microsoft and standardised by SMPTE. It is included in Windows Media. http://www.smpite.org/
VCH	Virtual Collision Handler	
VCL	Video Coding Layer	
VI	Video	
VO	Voice	
VOD	Video On Demand	An umbrella term for a wide set of technologies and companies whose common goal is to enable individuals to select videos from a central server for viewing on a television or computer screen. VoD can be used for entertainment (ordering movies transmitted digitally), education (viewing training videos), and videoconferencing (enhancing presentations with video clips). Although VoD is being used somewhat in all these areas, it is not yet widely implemented.
VoIP	Voice over Internet Protocol	Voice over Internet Protocol (also called VoIP, IP Telephony, Internet telephony, and Digital Phone) is the routing of voice conversations over the Internet or any other IP-based network. The voice data flows over a general-purpose packet-switched network, instead of traditional dedicated, circuit-switched voice transmission lines.
VoIPSA	Voice over IP Security Alliance	VoIPSA is an open, vendor-neutral organization, made up of VoIP and information security companies, organizations, and individuals. VoIPSA is a non-profit organization established in February 2005. VoIPSA formulates its mission to drive adoption of VoIP by promoting the current state of VoIP security research, VoIP security education and awareness, and free VoIP testing methodologies and tools. http://www.voipsa.org/
VoWLAN	Voice over WLAN	Voice over IP over a Wi-Fi network.
VPN	Virtual Private Network	A VPN is a private communications network usually used within a company, or by several different companies or organizations, to communicate over a public network. VPN message traffic is carried on public networking infrastructure (e.g. the Internet) using standard (often insecure) protocols, or over a service provider's network providing VPN service guarded by well defined Service Level Agreement (SLA) between the VPN customer and the VPN service provider.
VQEG	Video Quality Experts Group	The Video Quality Experts Group (VQEG) is a group of experts from various backgrounds and affiliations, including participants from several internationally recognized organizations, working in the field of video quality assessment. The group was formed in October 1997 at a meeting of video quality experts. The majority of participants are active in the International Telecommunication Union (ITU) and VQEG combines the expertise and resources found in several ITU Study Groups to work towards a common goal. http://www.its.bldrdoc.gov/vqeg/

VU	Visiting User	
WEP	Wired Equivalent Privacy	An implementation of RC4. It is part of the IEEE 802.11 standard (ratified in September 1999), and is a scheme used to secure wireless networks (Wi-Fi). WEP was designed to provide comparable confidentiality to a traditional wired network, hence the name. A new standard, IEEE 802.11i, provides improved security feature. See also WPA/WPA2. www.ieee802.org http://www.wifialliance.org
Wi-Fi	Wireless Fidelity	A term for certain types of wireless local area network (WLAN) that use specifications in the 802.11 family. The term Wi-Fi was created by an organization called the Wi-Fi Alliance, which oversees tests that certify product interoperability. A product that passes the alliance tests is given the label "Wi-Fi certified" (a registered trademark). http://www.wifialliance.org
WiMax/ WMAN	Wireless Metropolitan Area Network	Commonly referred to as WiMAX or less commonly as WirelessMAN™ or the Air Interface Standard IEEE 802.16. A specification for fixed broadband wireless metropolitan access networks (MANs) that use a point-to-multipoint architecture. Published on 8 April 2002, the standard defines the use of bandwidth between the licensed 10 GHz and 66 GHz and between the 2 GHz and 11 GHz (licensed and unlicensed) frequency ranges and defines a MAC layer that supports multiple physical layer specifications customized for the frequency band of use and their associated regulations. 802.16 supports very high bit rates in both uploading to and downloading from a base station up to a distance of 30 miles to handle such services as VoIP, IP connectivity and TDM voice and data. http://www.ieee802.org/16/ , http://www.wimaxforum.org/
WLAN	Wireless Local Area Network	This is a generic term covering a multitude of technologies providing local area networking via a radio link. Examples of WLAN technologies include Wi-Fi (Wireless Fidelity), 802.11b and 802.11a, HiperLAN, Bluetooth and IrDA (Infrared Data Association). A WLAN access point (AP) usually has a range of 20 –300 m. A WLAN may consist of several APs and may or may not be connected to Internet.
WMM	Wireless MultiMedia	A term used by the Wi-Fi Alliance.
WMV	Windows Media Video	
WPA	Wi-Fi Protected Access	An improved version of WEP (Wired Equivalent Privacy). It is a system to secure wireless (Wi-Fi) networks, created to patch the security of WEP. As a successor, WPA implements the majority of the IEEE 802.11i standard, and was intended as an intermediate measure to take the place of WEP while 802.11i was being prepared. http://www.ieee802.org , http://www.wifialliance.org
WPA2	Wi-Fi Protected Access 2	An extension to WPA that includes the remaining elements of IEEE 802.11i. http://www.ieee802.org , http://www.wifialliance.org
WWW	World Wide Web	An international, virtual network based information service composed of Internet host computers that provide on line information. A hypertext-based, distributed information system/service created by researchers at CERN in Geneva, Switzerland in 1991. Users may create, edit or browse hypertext documents. The clients and servers are freely available. http://www.w3c.org
xDSL	(Any) Digital Subscriber Line	Various configurations of digital subscriber line: X = ADSL – asymmetric, VDSL – very high speed, SHDSL – single pair high speed, SDSL – symmetric, HDSL – high speed. See DSL.