Voice-over-IP communication in wireless networks

Derk Wesemann, January 2004

Content:

1. Abstract

Voice-over-IP telephony has been implemented in many ways in wired networks, and is often used in companies for conferencing. As wireless networks have been developed in large numbers during the last years, these networks have become cheaper and more attractive to use, which implements new possibilities for mobile telephony.
In this area, voice-over-IP competes with other standards for mobile phones, like GSM, GPRS and UMTS. All these standards offer voice connections as well as data services, but there is a significant difference between these standards and VoIP-solutions: GSM, GPRS and UMTS have to be administrated by an external provider, who is responsible for the connectivity and billing. Especially for companies, it is attractive not to be dependant from an external organisation, but being able to care for their internal telephony and data services themselves. Taking a look at the costs for mobile telephony, this is a great advantage for all those employees which need to be equipped with a mobile phone, but work within the company's range most of the time, as an example.

This paper takes a look at current solutions for voice-over-IP protocols and compares their ability to fit in a wireless network like 802.11x, formally known as WLAN. It also addresses the basic problems of voice-over-IP connections in a wireless (mobile) environment, the necessity of the VoIP-connection to be able to handle changing parameters during a call, like changing addresses, routes, bandwidths, etc.

Current solutions for VoIP-protocols that will be addressed in the first part of this paper are the H.323-standard, developed by different workgroups of the ITU (International Telecommunications Union) and the SIP-protocol, which is a standard founded by the IETF (Internet Engineering Task Force). The second part will show how these protocols fit in a WLAN-environment. As examples for wireless networks, the 802.11x-series were chosen.

2. The session initiated protocol (SIP)

SIP is a application-layer control protocol that provides advanced signalling and control functionality for a large range of multimedia communications. It is used for establishing sessions in an IP network environment. A session could be a simple two-way telephone call or it could be a collaborative multi-media conference session. The ability to establish these sessions enables o lot of innovative services, ranging from instant messaging to video telephony. Over the last couple of years, SIP has become a standard protocol for signalling voive-over-IP-calls. SIP is an RFC standard (RFC 3261) from the Internet Engineering Task Force (IETF).

The abilities of the SI protocol are only to initiate, terminate and modify a session. It does not care about the session content, which makes it flexible in use and provides extensibility and adoptability to many scenarios. SIP has been developed to fit into the world wide web like HTTP and SMTP, offering telephony applications like any other internet service. For example, it uses the same URL-addressing scheme than HTTP.

SIP represents a modular approach to IP telephony protocols. It deals with generalized sessions. This involves finding potential call participants and contacting them as they move from place to place, changing their location, addresses and even the equipment they are using. Calls may require the use of multiple streams of various media, and very large numbers of participants might be involved in a call - and even joining and leaving in a constantly changing topology. This is supported by the ability of name mapping and redirection services, which allow the users to be available by one single identifier, which is consistent through changing topologies and networks.

A session itself is described on two distinct levels: The SIP-protocol takes care of addressing and protocol processing features, such as quality of service features ore signalling. The content of the exchanged media streams are defined by other protocols. The IETF suggests the Session Description Protocol (SDP, IETF RFC 2327) for describing the contents. SDP is, in fact, not a protocol but a structured, text-based media-description format that can be carried in the SIP message body. Since the message body is transparent to SIP any session description can be transferred. This makes this protocol flexible to use, for SIP sessions are not restricted to telephony calls or conference capabilities. The following figure no.1 takes a closer look at the architecture and how it fits in the ISO/OSI reference model:

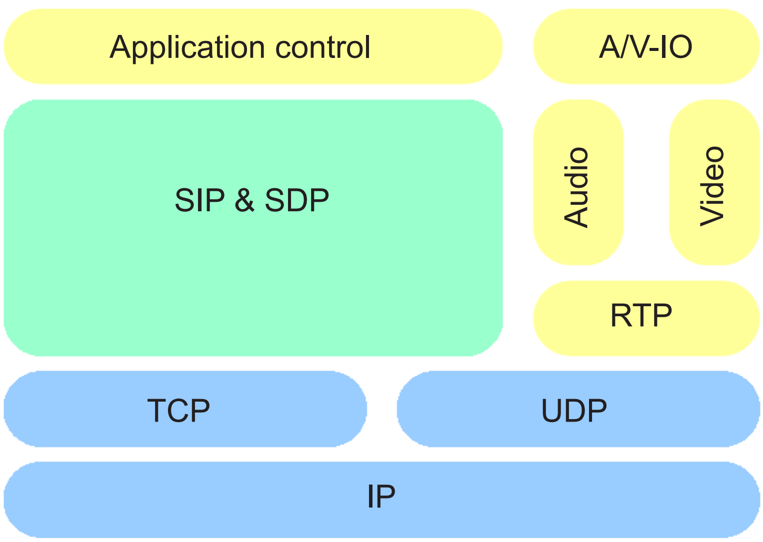

Figure no.1: SIP architecture

For establishing, maintaining and terminating a connection, the SIP-protocol supports the following five functionalities:

- User location: Determining a user's identifier
- User availability: Status modes, which show whether the user is willing to join or not
- User capabilities: Determining the media formats and parameters
- Session setup: Establishing session parameters on both caller and called side.
- Session management: Terminating or transfer sessions, modifying session parameters or services

One important thing to use when trying to establish an SIP-connection is the identifier of the user that should be called. This identifier is called SIP-URI (Uniform Resource Identifier). This SIP-URI is similar to an email-address, it typically consists of a user name and a domain name, as an example sip:derk@web.de. The call setup is initiated similarly to the HTTP request/response-model. An example is shown in figure no.2:

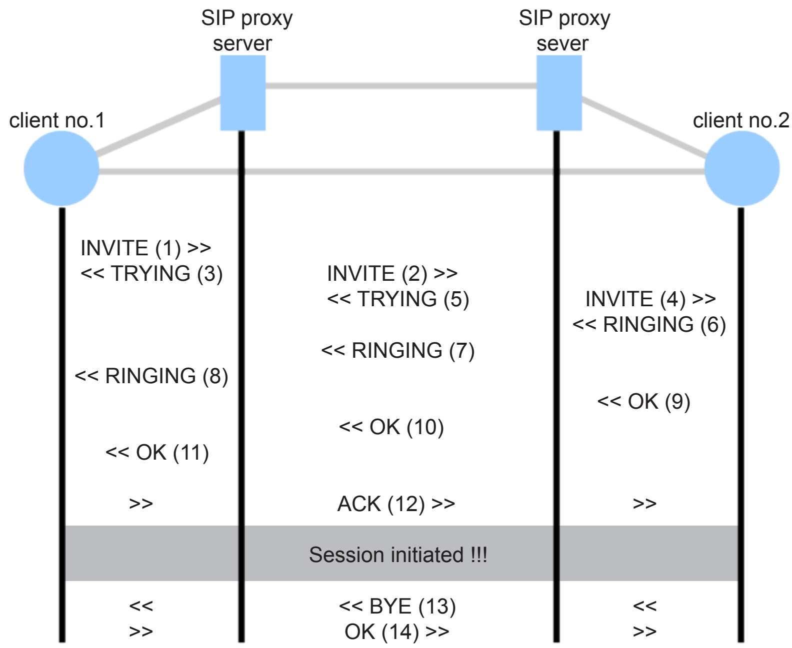

Figure no.2: SIP call setup

The SIP-connections is initiated by an INVITE-request. The caller does not need to know the location of the person that is called, because the request is sent to an SIP-server (proxy). This server resolves the domain name of the callee and forwards it to the according domain server, which is able to resolve the user name and contacts the according address. All call signalling messages are passed in text form, respectively UTF-8 encoded.

One drawback of the SIP-protocol becomes evident when assumpting the case of a moving client, the general case in a wireless network. The SIP-protocol does not directly provide functionality for managing the case of a moving client during a call, namely the handover from one network cell (access point) to another. The change of the device's network address would imply the call to be interrupted and re-initiated by an automatic INVITE-request.

3. The H.323-protocol

This protocol has been defined by the ITU (International Telecommunications Union) starting in 1995. In December 1996, the first version has been released, which was named to be a “standard for real-time video-conferencing over non-guaranteed quality of service LANs”. Talking about the H.323-standard nowadays referres to the fourth version. This standard is well defined system architecture, which implements guidelines for the entire call-setup, call control and the call’s media content. The H.323-protocol is currently the leading technology in VoIP-connectivity, offering voice- and video-conferencing over unreliable packet-based network connections. It does not provide quality-of-service-capabilities, although these service are often implemented in IP-networks nowadays.
The H.323-protocol was not designed completely new, but contains a vast number of other protocols (H.225.0-CC, H.320, H.324, H.450, H.332), and reuses other standards like the real-time protocol (RTP) and the real-time control protocol (RTCP). Some of it’s components were developed especially for this protocol, for example the protocol for registration, admission and status (RAS). That’s why the H.323-protocol is often called an “umbrella standard”. The different entities that make up an H.323 network include gateways, terminals, and Multipoint control units along with a gatekeeper. These four components are described a bit more in detail in the following:

Terminals:

Terminals represent the end points of a H.323 connection and can be realized in hardware or software. Each communication is initiated and ended by a terminal and can make use of several codecs for audio transmission. The codec that is used for a current connection has to be specified within the H.245-part. In a local area network (LAN), the terminals can communicate directly. For connections to other networks, an additional component is mandatory: The gateway.

Gateways:

Gateways are optional components in an H.323-environment. Gateways allow a conversion of different data formats in the transport or the control protocol, and also form an interface for connections outside the H.323-environment, for example PBX-terminals. Communication between terminals and gateways is done through H.245 and H.225.

Multipoint control units:

Multipoint control units (MCU’s) become necessary when a conference with more than two users is established. An MCU consists of a Multipoint Controller (MC) and a Multipoint Processor (MP). Both are used to mix and deploy the audio streams from all participants to their destinations.

Gatekeepers:

Gatekeepers serve various control and management functions within one H.323-zone. Only one gatekeeper per zone is permitted, and all terminals have to use the gatekeeper’s services. The main tasks of a gatekeeper are address conversion and bandwidth management. The last point, bandwidth management, makes this component attractive to use in wireless networks. The gatekeeper can assign a certain amount of the total bandwidth to H.323-connections, and rejects further connections when this limit is reached. The RAS-protocol is used to enable this functionality.

The information flow exchanged between the participants in a H.323-connection can be sorted into the following five categories:

- Audio (digitized and coded) voice data
- Video (digitized and coded) full-motion image communication
- Data (files, text documents or images)
- Communication control (exchange of supported functions, channel control, etc.)
- Controlling connections (connection setup, signalling, termination, etc.)

The H.225-model, part of the H.323-suite, is responsible for the information flow towards the LAN interface. It is based on the transportation layer (layer 4). An overview of the placement of the H.323-standard in relation to the ISO/OSI-model is shown in figure no.3:

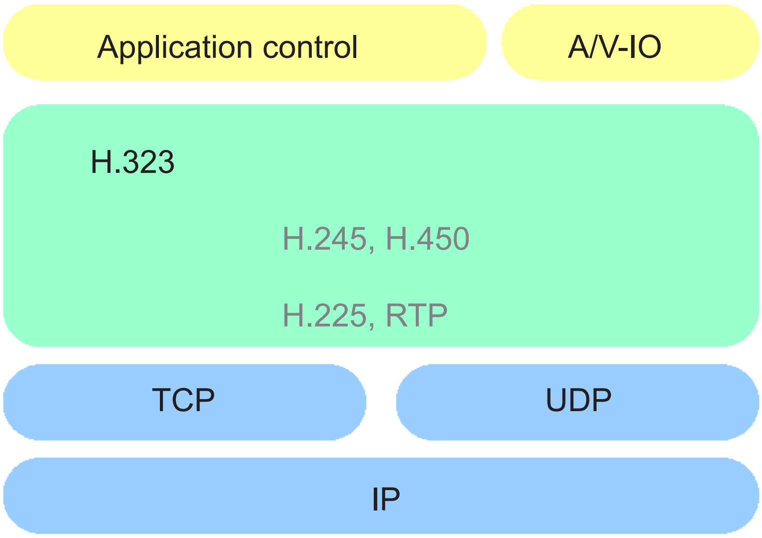

Figure no.3: H.323 architecture

H.323 v.4 defines the basic call control and signalling for setting up point-to-point and multipoint multimedia conferences. The RAS-protocol is responsible for signalling functions and call signalling functions. The difference between these two functions is that RAS signalling functions are required for end point registration, admission control and address resolution, while call-signaling functions are responsible for connection setup, capability exchange and logical channel procedures.

Taking a look at the architecture, H.323 includes services such as capability exchange, conference control, basic signalling, quality of service, registration, service discovery and further more. All data exchange in a H.323-environment is done bit-coded, in comparison to the SIP-protocol, where messages are UTF-8-encoded.

4. Comparison: H.323 vs SIP

The SIP-protocol and the H.323-standard follow quite different basic philosophies. The H.323 standards are more specified in the sense that they describe detailed protocols, state machines, and message flows for multimedia communication. This includes specific solutions for cross-sectional factors such as QoS, security, and mobility.
SIP, on the other hand, follows a more bottom-up approach, according to the IETF philosophy, where systems and applications are formed by combining generic modules. Every protocol standardized by the IETF should be independent of a specific application. Therefore, the SIP specification does not include issues such as QoS or mobility. SIP is a less complex protocol, that is only responsible to one function, namely setting up, leading and terminating sessions. All extended functionality is left over to other standards, that are interworking with the SIP-protocol.

There are also differences in the focus of H.323 and SIP in the past and at present. In the beginning, H.323 was concentrating on basic multimedia functionality, supplementary services, and interworking. Now that there exist sufficient solutions in those areas, one can observe a shift in focus to solution topics such as security, mobility, and QoS. SIP took another route, starting with the definition of a generic protocol to set up service sessions. Currently, SIP tends to put more focus on topics for specific applications, including supplementary services and interworking with legacy networks.

The choice between SIP and H.323 will at last be a question of the desired application. Additionally, these two protocols are the most discussed ones nowadays, but far not the only ones. Other currently developed approaches are MGCP/MeGaCo or Mobile IP, which does not directly care about voice connections, but allows roaming of mobile devices within local and wide area networks. For existing networks, the implementation of SIP is obvious to offer an easier integration than H.323, for no additional hardware is needed. The drawback, on the other hand, is that H.323 includes all necessary components for implementing additional (and sometimes needed) functionality, like quality-of-service or security issues. In comparison to the current state of SIP, those functions are added by making use of other protocols and standards, which leads to a higher overhead than what could be achieved with the H.323-standard. This behaviour is amplified by the fact, that in an SIP-network, messages are UTF-8-encoded, while H.323-networks exchange their data bit-coded.

All in all, the SIP-protocol is a bit more experimental than the H.323-standard. For H.323, several companies offer complete solutions, which results in better support at the cost of higher expenses. But since the SIP-development is pushed forward by companies like Cisco and Texas Instruments, and also has been implemented as a standard support in Windows XP, it is closing up on H.323, which is still leading in the area of VoIP-connectivity [4].

5. Wireless networks for VoIP-applications: The 802.11x standard

Talking about the 802.11-standard, often called WLAN of WiFi, it is necessary to distinguish between several sub-standards. The most recent versions and their specifications are illustrated in the following table:

standard	bandwidth	spectrum
802.11a	54 MBit/s	5 GHz
802.11b	11 MBit/s	2.4 GHz
802.11g	> 20 MBit/s	2,4 GHz

Table no.1: WLAN standards

The main difference between the a, g and b-standards is not only the bandwidth offered, but also the frequency range they are working on. This implies the disadvantage of incompatibility between the 802.11b- and 802.11a-standard, but also allows to create wireless networks that can cover a large area without interference at the boundaries. One functionality according to all named standards is the possibility to fall back to lower bandwidth values, when the connection quality decreases. As an example, in the 802.11b-standard, the bandwidth can be reduced stepwise from 11 MBit/s to 5.5, 2 or 1 MBit/s.

Taking a look at the bandwidth rates, these can only be achieved as peak values. The typical limitations are given by the MAC (media access control) -protocol that is used within the wireless network, which can be a CSMA/CA (carrier sense multiple access / collision avoidance) -protocol like DCF (distributed coordination function) or EDCF (enhanced distributed coordination function), or otherwise a TDMA (time division multiple access) -protocol like HCF (hybrid coordination function). DCF, EDCF and HCF are quality-of-service functions, that implement certain priorities to all connections. This leads to maximal throughput rates in wireless access points that are limited to about 37% of the available bandwidth for DCF and EDCF, respectively 75% for the HCF protocol. The following table no.2 shows the values for the different 802.11 standards, according to different modulation schemes:

Modulation	Throughput HCF	Throughput DCF/EDCF
54 MBit/s	~40 MBit/s (75%)	~20 MBit/s (37%)
22 MBit/s	~16 MBit/s (75%)	~8 MBit/s (37%)
11 MBit/s	~8 MBit/s (75%)	~4 MBit/s (37%)
5.5 MBit/s	~4 MBit/s (75%)	~2 MBit/s (37%)

Table no.2: Data throughput under non-optimal conditions

Concluding, this does not necessarily mean that using HCF will always lead to better results and should be preferred without reasoning. There are several pros and cons, addressing the choice of MAC-protocol on wireless networks, but it would lead to a too long writing to discuss these topics here. To keep it simple, the same assumptions that can be find in [1] were taken as a guideline. As a next step for evaluating the bandwidth issues of voice over IP in wireless networks, the following chapter addresses the impact of voice codecs on the network traffic.

6. Voice codecs for VoIP-communication: The G.7xx standard

A standard most often used for voice encoding in wireless networks is the G.711-standard with a sample rate of 64 kBit/s. Other standards are G.723, G.726, G.728 or G.729. Their basic differences are shown in the following table:

standard	bandwidth
G.711	64 kBit/s
G.723.1	6.4 kBit/s
G.726	32 kBit/s
G.728	16 kBit/s
G.729	8 kBit/s

Table no.3: Voice codec standards

The following illustration shows the typical appearance of an IP-packet carrying a voice-over-IP payload. The enclosed voice packets typically have a duration of 5, 10 or 20 ms sampled voice, depending on the codec that has been used.

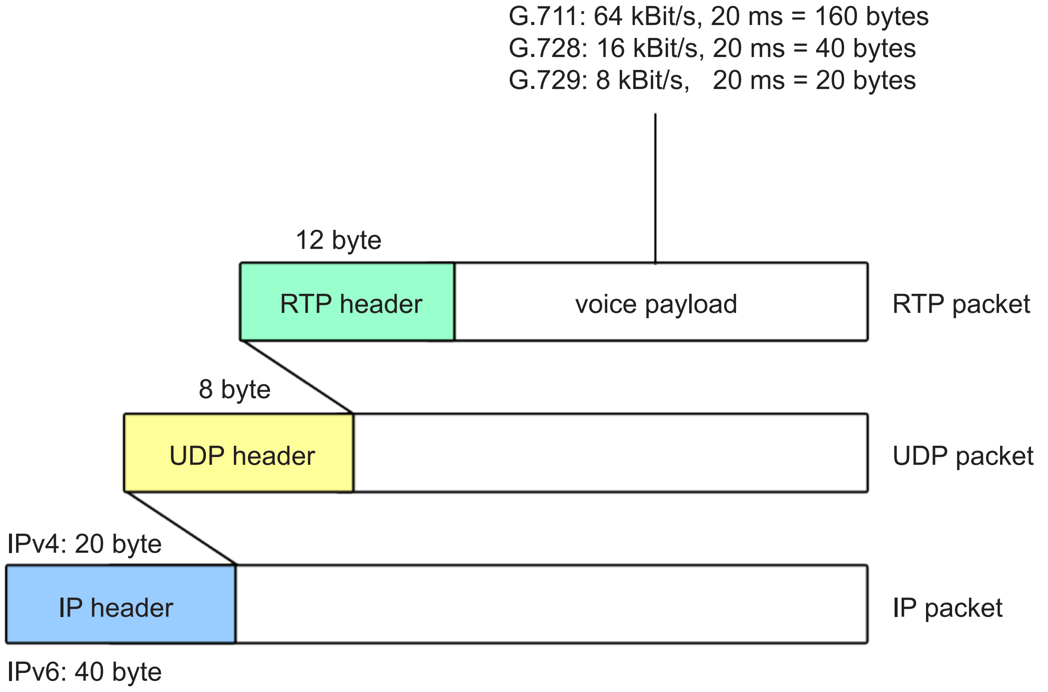

Figure no.4: IP packet appearance for VoIP-payload

We should note that a codec with a small bandwidth requirement does decrease the overall traffic, but as a drawback causes a high overhead. Taking the case of a 20 ms coded voice as an example, the overhead for an IPv4-packet with G.711-codec would be around 22%. In comparison, using the G.729-codec would lead to a packet efficiency of around 31%, which means that 69% of one packet are overhead and do not carry voice information.

7. Conclusion and future prospects

Finally, the evaluation of many different codecs, protocols and wireless network standards for voice-over-IP-functionality is a very up-to-date topic nowadays. This paper was designed to give an overview about the current approaches and trying to make a first step into the complex calculation of VoIP-requirements. The most recent and interesting calculations are those which address the needs for deciding which combination of protocol, codec and wireless standard fits best to a given infrastructure or environment, and what advantages and drawbacks should be concerned before making such a decision. The goal of all these researches, statistics and studies could be to give recommendations for all parameters, depending on the size of the network where voice-over-IP functionality shall be implemented. Another aspect could be not only to differ between several protocols, as an example, but to analyse which parts of a protocol are needed or can be modified or extended.

On the hardware side, one could imagine some kind of “holy grail” of mobility: A handheld device, that is capable of the most recent wireless network standards, not only 802.11x, but also bluetooth, GSM and GPRS, being able to switch between them when availability changes. Right now, the first mobile phones for the 802.11b-standard are manufactured, but they have a drawback of standby time, which is close to the first steps in GSM telephony.

Concluding, we can state that there is a great development speed in the area of VoIP-connectivity, leading to the often-used term “VoWLAN” when talking about mobile VoIP solutions. A major contribution of this development is the area of special IP protocols to support this functionality, and right there it is probably not hard to find a topic for further research.

7. References

[1] IEC online tutorial: VoIP over WLAN in a seamless Next-Generation Wireless Environment; IEC 2003

[2] Service Architecture in H.323 and SIP: A Comparison; J.Glasmann, W.Kellerer, H.Müller; IEEE 2003

[3] Internet Draft RFC2543 SIP:Session Initiation Protocol; J.Rosenberg, H.Schulzrinne, G.Camarillo, A.Johnston, J.Peterston, R.Sparks, M.Handley, E.Schooler; IETF February 2002

[4] H.323: The Leading Standard in Voice over IP; ITU-T 2003

Online material:

www.ietf.org
www.ieee.org
www.iec.org

and www.google.com ...