T. C. Schmidt, M. Wählisch, H. L. Cycon, M. Palkow, Mobility Support in Real-time Video Communication, in: B. Bodnar (Ed.), Proceedings of the Applied Telecommunication Symposium (ATS '03), The Society for Modeling and Simulation International, The Society for Modeling and Simulation International, California, 2003, pp. 72–77.

# Mobility Support in Real-time Video Communication [U]

**Thomas C. Schmidt[1], Matthias Wählisch[1], Hans L. Cycon[2] and Mark Palkow[2]**
**[1]Computer Center, [2] FB Ingenieurwissenschaften I**
**Fachhochschule für Technik und Wirtschaft Berlin**
**Treskowallee 8, 10318 Berlin, Germany**
**E-mail: {schmidt,mw,hcycon,mpalkow}@fhtw-berlin.de**

*Abstract:* **Mobile Internet devices recently gained much popularity and became widely available. Following an 802.11a/b or 3G standard they provide enough bandwidth to support data intensive communication services such as videoconferencing. Facing the emerging paradigm of ubiquitous computing and communication voice and videoconferencing over IP are on the spot to establish as standard Internet solutions. However, heavy infrastructural burdens and lacking ease of use as imposed by standards and missing consistent support of mobile serviceability hinder the acceptance.**

**This paper investigates basic aspects for supporting mobile conference users. We present a lightweight communication framework and conferencing software to overcome these deficiencies. A simple, ready-to-use global location scheme is proposed. Session mobility is investigated within the framework of MIPv6 with the special focus on temporal behavior.**

*Keywords:* **Peer-to-Peer Videoconferencing, User Locating, User Address Resolution, Mobile IPv6, Multicast Videoconferencing, Wavelet Transform**

## 1. INTRODUCTION

The designation of the Internet has changed. Networked devices, formerly placed at scientist and business desks, are now consumer parts and serve for information, communication and entertainment. Their dedication follows trends as clearly lie in mobility and ubiquitous computing at present. The vision of roaming users at roaming devices performing synchronous communication such as voice or videoconferencing over IP (VoIP/VCoIP) is around and rises new challenges for the Internet infrastructure.

Videoconferencing until now did not pick up as a popular way of communication like talking on the telephone or exchanging emails. This has several reasons. On the one hand traditional ISDN based systems are costly. On the other hand videoconferencing over IP still seems to be something for computer geeks which are happy with tiny moving heads and poor quality audio. Thirdly, no easily accessible framework for session based communication has been offered to the Internet community.

So, there are still large gaps between "talking stamps", heavily equipped video rooms and a seamless application in everyday life. But this gap is now narrowing. Videoconferencing is going to emerge into a lightweight day-to-day application. One reason is that more and more bandwidth becomes available to the public for reasonable prices.

There is also some remarkable progress in video/audio compression algorithms which compress the video data stream to less than 1% on the sending site and reconstruct it again to a high quality video sequence on the receiver site. A third reason is that more and more desktop videoconferencing software running on ordinary desktop PCs or laptops using internet connections are available. The video/audio quality and the easy to use user interface combined with all kinds of application sharing will soon convince more and more users to enrich their communication habits by video components.

As broadband spreads on earth, wireless internet conquers the air. The availability of new, truly mobile IP enabled subnetwork layer not only carries along the requirement for dynamically locating users and for preserving communication sessions beyond IP subnets, but re-raises questions concerning quality of IP-service: The constant bit rate scenarios of voice and videoconferencing will appear significantly disturbed by packet loss intervals or delays exceeding 100 ms. In addition jitter needs to stay bound well below this limit. Thus, when heading towards VCoIP as a standard Internet service important steps for global usability focusing on ease and quality have to be taken.

Addressability issues of mobile devices as well as for VCoIP participants raises the demand for the maturing next generation Internet protocol. IPv6, about ready for rollout today, incorporates an interesting collection of additional functionality such as multicast availability, QoS routing, mobility and security. Many of these features appear significantly valuable for a videoconference solution and an early orientation towards IPv6 in the context of VCoIP advisable.

[U] This work was supported in part by the EFRE program of the European Commission and by the German Bundesministerium für Bildung und Forschung.

In the present paper we address the issue of global, decentralized VCoIP architecture and solutions in section 2. Furthermore the paper is organized as follows. In section 3 we introduce the idea of a distributed framework for mobile user location. Section 4 presents concepts and experiments to IPv6 session mobility. Finally, section 5 is dedicated to conclusions and an outlook on future experiments and developments.

## 2. VCoIP ARCHITECTURES AND SOLUTIONS

### 2.1 Related Works

Video conferencing over IP still waits to be established as a regular communication service. To progress its dissemination throughout the Internet community the most simple application scenario should be kept in mind: Any Internet user may call any online partner by just starting an appropriate software tool and addressing a common name.

Up until now the employment of video conference applications has been dominated by ISDN systems. This traditional technology offers a person to person, respectively meeting oriented, private service as telephony in general does. The communication paradigm consists of a point to point connection between dedicated devices under specific user attendance.

VCoIP is in contrast embedded into general Internet-connected working devices and today oriented towards more or less public conference groups. As employment of VCoIP grows more mature, though, the need for meeting oriented, private sessions urgently has to be met. Since it addresses people instead of devices it should adapt to the common internetworking communication paradigm of mobile users accessing services, not equipment.

The traditional, ISDN compatible architecture of VCoIP systems has been defined in the ITU standard H.323 [1]. Central parts of this model are derived from a client-server principle with a Multipoint Control Unit (MCU) serving video streams in multipoint conferences and a Gatekeeper providing connection control and address translation. One advantage of the MCU facility design lies in its ability to transform data streams between different video/audio codecs. The major disadvantages of course are drawn from the request for heavy infrastructural changes and significant latency additives.

The H.323 architecture must be considered as local in the sense that all participants need to agree on the common MCU and Gatekeeper servers which, obviously at least for the MCU, suffer from severe scaling deficiencies. No global naming is defined except for telephone numbers handled by ISDN gateways and the Q.931-compatible signalling protocol H.225.0. H.323 concepts centre around the ideas of telephone-based wide area connectivity and are obsolete by the simple observation that use of video conferencing via telephony does not grow. Consequently attempts are made to overcome local restrictions in addressing by interconnecting Gatekeepers via meta-directory servers as done in the Video Development Initiative [3].

H.323 terminals may be used independent of servers for bilateral conferences. In this way MS NetMeeting a.o. operate. The serverless extension to multipoint abilities in the IP world are most efficiently done via multicast transport, where any client in the conference simultaneously takes the role of multicast source and destination. Multicasting is employed at the price of communicating in more or less full public. Multicast features do not conform to H.323 and have been implemented e.g. by the Mbone Tools [4], Vcon [5], Ivisit [6] and ISABEL [7].

A fairly general attempt to overcome H.323 has been made with the Session Initialisation Protocol (SIP) [8]. SIP covers negotiations about user capabilities, user availability, user localisation, the call set-up by Session Description Protocol and the handling of the calls itself. SIP introduces its own infrastructure of servers which actively communicate by using SIP-URLs or other network protocols such as ICMP. SIP is open to store persistent information in common databases such as LDAP directories, but adheres its own server communication layer.

The SIP concept proposes either a significant roll-out of SIP self-learning, interrelated infrastructure or just the presence of single, isolated information servers. In the latter case, strategies to locate these information servers remain vague. Both SIP and H.323 have the drawback of exchanging addresses within the protocol payload and are thereby severely hindered in NAT traversal and need dedicated adaptation to IPv6.

### 2.2 The daViKo Videoconferencing System

The digital audio-visual conferencing system daViKo [2] forms a serverless multipoint video conferencing software without using a MCU developed by the authors (see fig. 1). It has been designed in a peer-to-peer model as a lightweight Internet conferencing tool aimed at email level use. Guided by the latter principle, daViKo refrained from implementing H.323 client requirements [1].

The system is built instead upon a fast, highly efficient video codec, based on the wavelet algorithm described below. By controlling the coding parameters appropriately, the software permits scaling in bandwidths from 64 to 4000 kbit/s on the fly. Audio data is compressed using an MP3 algorithm with latencies below 120 ms depending on the chosen buffer size. Audio and video streams can be transmitted as unicast as well as multicast. An application-sharing facility is included for collaboration and teleteaching.

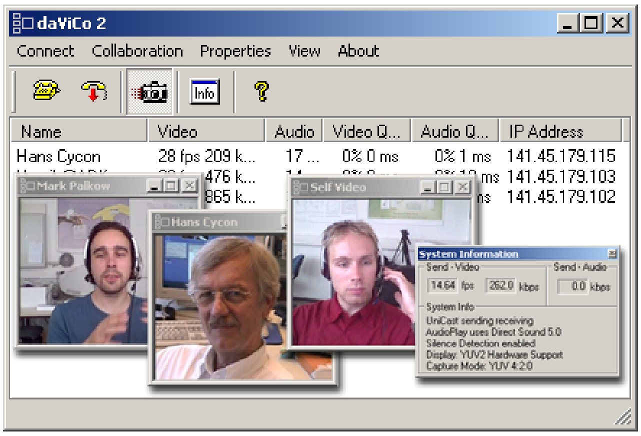


**Figure 1: The daViKo Conferencing Tool**

Due to low bandwidth requirements, daViKo is well suited to long distance video-conferences on a best effort basis. To strengthen its global usability even on mobile devices, the user location scheme described below has been implemented into the system as well as advanced IPv6 network capabilities.

### 2.3 Wavelet-Based Real-Time Video Codec

Our coding scheme is basically a transform coder together with a simple frame-based temporal prediction loop as shown in figure 2. The transform coder consists of the reversible discrete wavelet transform (DWT) which decorrelates the signal, a quantizer (Quant.), and a lossless precoding/entropy coding step which compacts the data produced by the quantizer. For exploiting the temporal redundancy in a video sequence, only the residual signal between the current frame and the previous reconstructed one will be coded in the transform coding step, as shown in figure 2. Due to the linearity of wavelet transforms we compute the residual frame in the wavelet transform domain instead of calculating the frame difference in the spatial domain. This saves an inverse transform step in the temporal prediction loop.

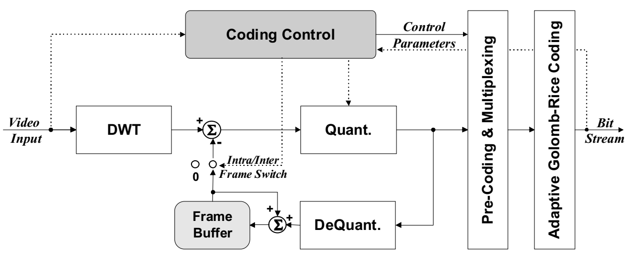


**Figure 2: Block diagram of our video coding scheme.**

Since we do not employ motion estimation techniques a periodic intra frame fresh-up after 99 inter, i.e., differentially coded frames is installed in order to avoid a too fast fading of image quality. However, this inter-to-intra frame relation can be manually chosen depending on the dynamical character of the video sources to be encoded.

In order to guarantee a constant transmitted bit rate on the average, we have implemented a backward operating coding control mechanism. We use a ring buffer with a dynamic level control which measures the quantized data stream. The bit rate can be adjusted on the fly by using control parameters during encoding. Note that these parameter changes have to be signalled to the decoder and therefore are integrated into the encoded bit stream.

Note also that this codec is a stripped-down speed-up version derived from conceptual ideas presented in [10] and [11].

**Results**

In native implementations, the video codec encodes and decodes 25 CIF frames (352 x 288 pixels) simultaneously on a 500 MHz Pentium machine. Alternatively, 5 frames in PAL (720 x 576) resolution may be processed, where frame rate is expected to increase with forthcoming algorithmic improvements. The image quality is better or comparable with MPEG 4 / H.263 Coders. At moderate motion complexity, this frame rate produces a bit rate of ca 200 kb/s while sustaining very good visual quality.

The codec has also been ported to JAVA as part of a Web streaming system [12]. The JAVA codec running in an applet still decodes or encodes 5 CIF frames per second in real-time or, more appropriately, QCIF format with 25 frames.

## 3. A LOCATION SCHEME FOR MOBILE USERS

Videoconferencing is a synchronous form of communication requesting online presence of the participants. To retrieve the information on how to direct data flows to the appropriate user, his current device address needs to be resolved somehow. As device addresses change with mobility and as users may move between devices, a static address selection or any out of band information on user's presence are inappropriate.

Instead a dynamic user session recording has proven advantageous. In the system introduced here, we denote this by a User Session Locator (USL) and store appropriate session information in an LDAP directory server. The videoconference clients update information about ongoing sessions regularly, so that outdated session records can be identified by their timestamps. The USL server can be arranged within a local infrastructure not only to enhance scalability by distribution, but also to adopt local knowledge of the identity of users as well as a method for authentication. Note the importance of authentication procedures for user session registration: private communication channels are directed by advertising user session data. Also, authenticated user session data may serve as a weak mechanism of identification: a callee may verify an agreement of IP and the user address of the caller by searching the USL session registration in a trusted domain.

Whereas a local search on a USL server can be performed in a straight forward fashion (see fig 3), the global user look-up problem is reduced to deciding on unique user addressing and discovering the appropriate directory server for a given address. Current solutions either concentrate on a centralized directory as does MS NetMeeting with the MS Internet Locator Server [15] or perform an Internet wide user based routing as is the purpose of the SIP server infrastructure. Since SIP does not prescribe a specific addressing scheme, it needs to cope with user addresses changing under mobility.

Currently, the only uniformly available user addressing scheme on the Internet is given by mail. Mail addresses are not only globally unique but also device independent, commonly known or easily retrieved. Several vendors have noticed the uniqueness and popularity of mail naming, so that calling a videoconference user by his mail name has gained some popularity. Our system *restricts* user addressing to mail addresses because of its convenience and ease of use. In adopting this restriction we radically break with telephone compatibility.

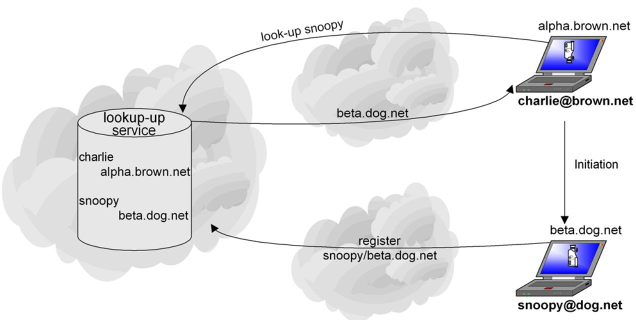


**Figure 3: Centralised User Look-Up**

In addition the Internet mail system provides a mechanism for resolving user location through its interaction with the Domain Name System via the MX record type for referencing a mail exchanger. Employing this commonly available Internet infrastructure we chose a simple strategy for locating a user's session directory. DNS data provided today are ready to cope with it: because the mail exchange record indicates a physically present domain where any requested user is identifiable along with a method of authentication, it is the appropriate location for a USL server. Within this domain, the look-up server can be identified by the common approach of a naming convention, i.e. usl.<mailexchanger-domain> [14]. Consequently, a global user look-up proceeds in two steps. Firstly, the MX record for the target user is requested, and secondly, the directory server hostname formed from the above naming convention is resolved (see fig. 4).

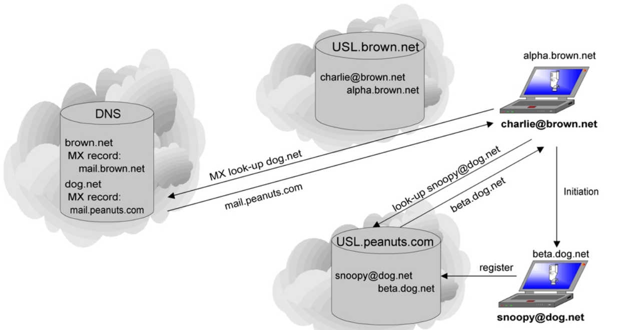


**Figure 4: Distributed User Location Scheme**

Though simple, this user session information architecture neither relies on infrastructural changes nor requires dedicated user knowledge on the application side. From the mobility point of view the USL servers play the role of distributed user home agents. Note that in contrast to H.323 gatekeepers or SIP servers the USL server consists of a passive session record store and can be realised by an unmodified standard LDAP server such as OpenLDAP. It is easily integrated into existing local infrastructure and may establish videoconferencing as a serious, regular Internet communication service. For more detailed reading we refer to [13].

## 4. SESSION MOBILITY UNDER IPv6

### 4.1 Changing Networks

Viewing roaming users carrying mobile Internet devices we have to consider changes of IP subnets during conferencing sessions. This could not only be relevant in fast moving cars or trains, but may also happen when a moving device is forced to alter its provider. As it is common standard in mobile phones, a strong concern of preserving VCoIP session beyond renumbering should be stated. The next generation Internet grants the ability of coping with multiple and changing addresses, thus giving rise to devices migrating their connectivity while running in service.

The fundamental approach to Internet mobility is the Mobile IPv6 (MIPv6) Internet-Draft [16]. MIPv6 transparently operates address changes on the IP layer as a device moves from one network to the other by sustaining initial IP addresses and hiding the different routes to the socket layer. In this way hosts are enabled to maintain transport and higher-layer connections when they change locations. Since MIPv6 permits a mobile node to join or remain in a multicast group by use of its care-of address or by tunnelling through its home agent, multicast transparency is preserved, as well.

An alternate approach to application persistency under mobility is grounded on the Stream Control Transmission Protocol (SCTP). Initially designed for network redundancy SCTP allows for multihoming of a single socket. The addip proposition [17] of extending this functionality to adding and deleting IP addresses gives rise to an address handover on the application side. Even though address renumbering on the transport layer is placed improperly, mobile SCTP [18] carries the justification of performing a rapid handover on the client side, only, without any provisions in the layer 3 infrastructure. Mobile SCTP, though, conflicts with single bound layer 2 protocols s. a. 802.11, does not support MSCTP node discovery and there are strong arguments for gaining transport mobility as a combination of SCTP and MIPv6.

### 4.2 VCoIP over MIPv6

Real-time video communication imposes stern quality of service requirements on the underlying network infrastructure. 100 ms real-time carry relevant information, a spoken syllable for instance in the audio case. More generally network disturbances of more than 300 ms interrupt a video conference at the user's level, whereas perturbations lasting less than 100 ms remain tolerable and may even be adjusted by jitter-hiding buffers [9]. Thus timing of handover procedures in MIPv6 forms a critical issue.

In entering a new network, i.e. completion of the layer 2 (L2) handoff, the MN instantaneously has to perform an automatic address reconfiguration followed by binding updates with its home agent (HA) and the correspondent node (CN). During this handover procedure the mobile node is unavailable until the HA has learned its new care-of address. Packets may than proceed through the HA as a forwarder with the likely result of increased delay and jitter.

The temporal performance of network changeover is determined by two different mechanisms: The link local procedures of L2 handoff with succeeding readdressing on the one hand, which only depend on local subnet topology, and the distant updates with HA and CN dominated by network geometry. Approaches to decrease the latter dependency are stated in [19] and [20] and include dedicated provisions within the network infrastructure.

### 4.3 Experimental Setup

Mobile Internet scenarios are quite unsettled and a large variety of phenomena still needs a careful analysis. In our current experiments we focus on the local handoff procedures in MIPv6 over 802.11b wireless LAN following the setup of fig. 5: Starting from its home network a mobile node (MN) proceeds through two 802.11b wireless LANs.

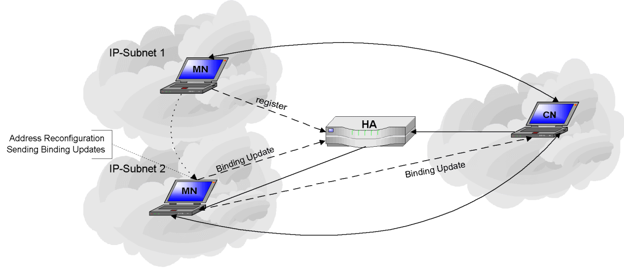


**Figure 5: Mobile Scenario**

Our testbed consists of a Debian Linux 2.4.19, MIPL mobile IPv6 0.9.4 machine as a home agent, a router based on FreeBSD 4.6-STABLE with rtadvd and Win2000 clients with MSR (1.4) TCP-IPv6 driver 5.0.21955.1620 as well as Debian Linux 2.4.19 with MIPL 0.9.4 as mobile and correspondent node. Stacks of the *mobile* nodes and the router advertisement daemon were modified with respect to temporal and MIP-signaling (DAD) behavior. The mobile nodes send and receive numbered and time-stamped UDP packets following a predefined trigger of typically 10 to 20 ms. These packets are reflected by the correspondent node. All events were packet wise recorded using a network sniffer.

### 4.3 Results

Packet loss, packet roundtrip time and jitter occurrences are shown in figure 6 as averages over events of the Linux MN's traversal. With a router advertisement interval (MinDelayBetweenRAs) of 50 ms 90 % of the events exhibit an interval of network disturbance below or equal to 100 ms.

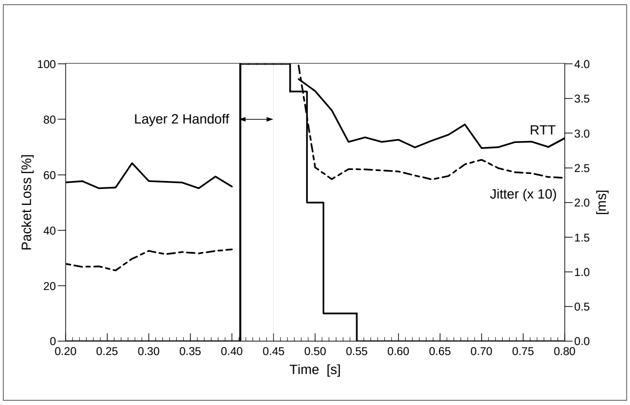


**Figure 6: Empirical Results**

The handover effect is dominated by packet loss which is due to L2 handoff (~40 ms) and MIPv6 updates (10 – 60 ms). The link local IP readdressing hereby attained an average duration of 25 ms. Note that the Linux MIP stack discovers a change of network only through layer 3 router advertisements, which in our scenario, still conformal to [16], produce a base load of 0,5 % 802.11b network capacity.

The more efficient algorithm for a MN lies in discovering the network change through the local L2 stack and actively performing a router solicitation call. Even though there is no standard for signaling 802.11 L2 events to IP, the Windows MIP stack is aware of L2 handoff by means of its proprietary NDIS architecture. An reduction of the timer variables MAX_RA_DELAY_TIME and MAX_RTR_SOLICITATION_DELAY at the router and the MN may then reduce MIP local readdressing time well below 5 ms without adding base load to the network.

In a forthcoming paper [9] we will report on more detailed experimental results.

## 5. CONCLUSION AND OUTLOOK

Video Conferencing over IP exhibits an enormous potential to become a regular standard service in an upcoming world of ubiquitous mobile Internet devices. However, the distribution of VCoIP presently is retarded because common approaches rely on significant changes of Internet infrastructure. We presented a proposition, both communication framework and conferencing software, to overcome these obstacles from the lightweight side.

Future development of our system will more closely evolve according to standards. The advancement of our video codec will be part of the ITU-T standard H.264 resp. the MPEG standard 'Advanced Video Codec' (AVC).

Significant challenges are embodied in the Internet mobility tasks. Even though a principle feasibility of IP mobility under real-time video communication could be demonstrated, MIPv6 handover procedures need tightening. Future improvements need to focus on a reduction of packet loss probabilities as seem attainable by reviewing MN stack properties and appropriate buffering opportunities.

## ACKNOWLEDGEMENTS

We like to express our deep gratitude to Stefan Zech, who had his helpful hands in many technical snarls and always shared an open ear or three.

## REFERENCES

[1] ITU-T Recommendation H.323: "Infrastructure of audio-visual services – Systems and terminal equipment for audio-visual services: Packet-based multimedia communications systems", Draft Version 4, 2000.

[2] "The daViKo homepage", http://www.daViKo.com, 2002.

[3] "Video Development Initiative homepage", http://www.vide.net, 2002.

[4] Mbone tool download, ftp://ftp.ee.lbl.gov/conferen ing/, 1996.

[5] "The VCON homepage", http://www.vcon.com, 2002.

[6] "The IVISIT homepage", http://www.ivisit.com, 2002.

[7] "The ISABEL Homepage", http://isabel.dit.upm.es, 2002.

[8] J. Rosenberg et. al.: "SIP: Session Initiation Protocol", RFC3261, June2002.

[9] T.C. Schmidt, M. Wählisch, H.L. Cycon, M. Palkow: "Mobility Aspects in IPv6 Videoconferencing", in preparation.

[10] D. Marpe, H.L. Cycon: "Efficient Pre-Coding Techniques for Wavelet-Based Image Compression", 1997, Proc. PCS '97, pp. 45–50.

[11] D. Marpe, H.L. Cycon: Very Low Bit-Rate Video Coding Using Wavelet-Based Tech-niques, IEEE Trans. on Circ. and Sys. for Video Techn., 1999, 9 (1), pp. 85–94.

[12] B. Feustel, T.C. Schmidt: "Media Objects in Time --- A Multimedia Streaming System", Computer Networks, 37/6, November 2001, pp 727--735, Amsterdam 2001.

[13] T.C. Schmidt, M. Wählisch, H.L. Cycon, M. Palkow: "Global Serverless Videoconferencing over IP", to appear in Future Generation Computer Systems, Elsevier, 2003.

[14] M. Hamilton, R. Wright: "Use of DNS Aliases for Network Services", RFC2219, October 1997.

[15] NetMeeting Resource Kit Contents, Chapter 3, "Finding People" http://www.microsoft.com/Windows/Net Meeting/Corp/ResKit/Chapter3/default.asp, 2002.

[16] D.B. Johnson et al: "Mobility Support in IPv6", IETF Internet Draft, Rel. 19, 29 Oct. 2002, work in progress.

[17] R.R. Stewart et al: Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration", IETF Internet Draft, Rel. 6, September 2002, work in progress.

[18] M. Riegel, M. Tuexen: "Mobile SCTP", IETF Internet Draft, Rel. 1, 9 August 2002, work in progress.

[19] H. Soliman et al.: "Hierarchical Mobile Ipv6 mobility management", IETF Internet Draft, Rel. 7, October 2002, work in progress.

[20] R. Koodli (Ed.): "Fast Handovers for Mobile IPv6", IETF Internet Draft, Rel. 5, 30 Sept. 2002, work in progress.