Technical University Berlin

Telecommunication Networks Group

Current developments and trends in handover design for ALL-IP wireless networks

A. Festag, H. Karl, G. Schäfer
festag,karl,schaefer@ft.ee.tu-berlin.de

Berlin, 08/18/2000, Version 1.3

TKN Technical Report TKN-00-007

TKN Technical Reports Series

Editor: Prof. Dr.-Ing. Adam Wolisz

Abstract

This document investigates handover in *All-IP* wireless networks. It works out general trends in network services and architectures of future IP-based wireless networks with an impact on handover design. The main contribution of this document is the description and comparison of the handover approaches which have been identified as state-of-the-art in the research community: IETF Mobile IPv4 [17], Extensions of IETF Mobile IP (hierarchical foreign agents [7, 8, 9], route optimization [18], flexible mobility support [33]), IETF Mobile IPv6 [11], Reverse Address Translation (RAT) [24], multicast-based handover [6, 14, 26], HAWAII [19, 20], Cellular IP [28], Mobile People Architecture [13], ICEBERG [29], Extended SIP Mobility [30]. There are two main results of the document: Foremost it is explained how the approaches solve the general mobility problem. Secondly it is stressed how the approaches meet the demands arising from new user and network requirements and technical opportunities of new technologies. The document concludes with requirements on the design of handover schemes for future IP-based wireless networks. [1]

[1] This work has been supported by a grant from Siemens.

Contents

List of Figures

Chapter 1

Introduction

Handover describes a mechanism in cellular networks that transfers the association of a mobile end system from one base station - which is presently active - to a new base station. In general handover is applied when a user moves through the coverage area of a cellular network and crosses cell boundaries. The handover between wireless cells of the same type (in terms of coverage, data rate and mobility) is often referred to as horizontal handover, whereas the handover between wireless cells of different type is characterized as vertical handover. Traditional IP-based mobility approaches, such as IETF Mobile IP, were designed with respect to horizontal handover. Thus, the vertical handover and other new services and network architectures pose new requirements on handover design.

Nevertheless the fundamental mobility problem in IP based networks still remains: IP protocols were designed for stationary end systems. The IP address of an end system identifies a host uniquely and also the IP subnet to which the host is attached. Therefore the meaning of the IP address is twofold: end point identification and location identification. When a host changes its point of attachment the IP address must be modified in order to route packets to the mobile's new subnetwork. Unfortunately, ongoing TCP connections break since the IP address is part of the TCP connection identifier and used at TCP connection setup.

Handover has received considerable attention in recent years. Foremost a number of system-specific solutions have been developed for GSM, GPRS and UMTS networks, for mobile extensions of ATM networks, as well as for wireless LANs, such as IEEE 802.11 networks. From the Internet point of view these solutions can be regarded as layer-2 solutions for wireless access networks working transparently to the IP layer. In addition transport layer approaches have been investigated, such as [2, 3, 12]. Finally, application-layer mobility support is common in typical client-server applications, such as email (POP), news and ftp (”smart download” with restart capability).

This report addresses *All-IP* wireless networks, which are based on Internet principles, services and protocols. The demand for *All-IP* networks have a multitude of consequences and open questions, e.g. is there an end-to-end IP concept between mobile hosts and corresponding hosts. One of the main question related to mobility support is for the requirements on the backbone elements in order to support such an All-IP concept. We envision a future generation wireless network as a pure IP-based network. Base stations and routers in the access network are IP addressable entities. We believe that mobility should be handled at network layer in cooperation with link layer and application layer.

The report presents a survey of handover strategies considered in recent research-oriented handover approaches. Handover is described as part of the overall mobility concept. Although the handover is stressed it is necessary to explain the basic mobility aspects.

The report is organized as follows: At first, basic usage scenarios are introduced. Then recent trends in network services and architecture in general and their impact on handover design are described in Chapter 3. Important research-oriented handover approaches are introduced in Chapter 4: It is intended to elaborate the salient points of current trends in handover design. Therefore, the basic approach and assumptions behind the schemes are stressed. The use of fundamental mechanisms and building blocks in each of these approaches are summarized in Chapter 5.

Chapter 2

Scenarios

First three different scenarios for the usage of wireless access in general are introduced [32]:

In the first scenario, which is referred to as *basic wireless access* the motivation for using wireless technologies is mainly avoiding installation of cable, which might be kind of cumbersome. In this scenario terminals can be moved exclusively within the range of a single, always the same access point, and the movements are slow (if any). On the other hand, even using the terminal at home gives good reasons for wireless transmission. One could consider this variant as generalization of cordless telephony. The major challenge in this case is assuring the proper quality of service to the terminal. Numerous wireless transmission technologies are already available, or will soon become available, for the support of such scenario: Wireless LANs, Bluetooth, IrDA belonging to the most frequently mentioned ones.

In the second scenario, which we will refer to as *nomadic wireless access* the terminal is expected to be moved over distances essentially exceeding the transmission range of a single access point. It is assumed, that multiple access points will be deployed over some area (which might be a building, a campus, a city or a continent) creating islands of connectivity around these access points. It is assumed for this scenario that terminals may switch between the access points only between the consecutive sessions. This movement takes usually times which are long as compared to the session duration (one might think here in the terms of a scientist visiting in turn several universities). In fact there are no hints in which location - close to which access point - the terminal might appear after movement. Let us stress that multiple parameters like supported bit-rate, error rate, the maximum speed of mobility and the supported range around the given access point are in general not identical even among individual access points supporting the same technology (due to static or dynamic set-up differences). Assuring a simple set-up in the new environment seems to be a major challenge for this scenario (different access points in distant locations might even support heterogeneous technologies). An additional challenge is assuring reachability under the original address in the actual, temporary environment as well as security considerations. This problem is frequently referred to as roaming.

Finally in the third scenario, which we will refer to as *true mobile access*, dynamic changes of the supporting access point during a session, usually referred to as handover, are expected to appear (possibly even several times during a single session). For mobile access to be attractive, the deployment of the access points should be more dense as for the nomadic wireless access, usually a significant (although not necessarily all) part of the area will be in the range of at least one of these access points. We frequently use the notion of coverage while referring to the ratio of the area being within the range at least one access point to the whole area under consideration. In this scenario. The grade of service continuity in spite of handover is one of the essential quality features for this scenario. Continuity of service might be expressed in terms of no information loss during the handover, sometimes even so called seamless handover, i.e. handover not observable by the user at all, is required. This requirement might result in a necessity for the terminal to remain during the handover in the range of both participating access points. In any case the requirement for frequent, possibly interruption-less handover implies usually a homogeneous system concept in which all the access points (and the end-system) are incorporated. In addition it seems almost natural (although not necessary!) to assume that also the transmission techniques will be always unified. In fact this is the case for the majority of solutions deployed or considered today, like GSM, GPRS or the emerging UMTS. The transmission range might be, technology depend, short or large. By the way: wireless access is attractive also in the case of no movement at all- this case is in the classical telephony referred to as the WLL- wireless local loop. Due to fixed positioning of the terminals relative to the access point several techniques for improvement of the quality of signal might be applied. We will not discuss this case in depth in this paper. It should be pointed out that the principle of nomadic access is in general NOT necessarily to be discussed in the context of wireless communication, nomadic computing can be also considered using wired transmission. It seems, however, that wireless transmission might encourage broader deployment of nomadic computing- think for example in the terms of passengers in airports or scientific conference participants.

Chapter 3

General trends in network services and architecture

Wireless communication and Internet have found pervasive deployment in the 1990ies. The technological progress was driven by developments in computing and communication devices, wireless and wire-line transmission technologies, communication protocols and standardization efforts. The long term evolution is sometimes referred to as *Ubiquitous Computing* (URL: http://www.ubiquitous.com):

> *"The most profound technologies are those that disappear.*
> *They weave themselves into the fabric of everyday life until they are indistinguishable from it."*

"The Computer for the 21st Century", M. Weiser [31]

The vision of ubiquitous computing is to enhance the computer use by making many computers available to anyone throughout the physical environment, by making them effectively invisible to the user. Transferring the picture to communication it means *communicating anywhere, anytime with anyone*. Having the vision of *ubiquitous computing and communication* in mind, the identification of short term trends in network services and architectures offer a framework for handover design. Such a framework has to be put into perspective with probable user requirements for future mobile communication systems as well as with opportunities arising from new technologies (where such technological developments often spark user requirements).

3.1 User and network requirements and technical opportunities of new technologies

Ubiquitous network access. Network access will be offered anytime with a basic quality of service. It is provided by full spatial coverage via a cellular network. Also, if more than one communication device with persistent network access is available, it is the users choice which device to use.

User mobility. While recent approaches for mobility in the Internet focus on end system mobility, future networks will also support user mobility. In that case a user is identified with an identifier, which introduces a new addressing concept into the Internet. For mobility support this requires mainly to locate people instead of end systems, to establish a sessions to a user which has assigned multiple temporary end system addresses. Finally sessions between users have to be maintained though a user changes the communication device while a session is ongoing. The concept of user mobility is not an inherent Internet concept, whereas it is common in e.g. second generation mobile cellular networks. Today's users of communication services have several identifications: email address, mobile phone number, telephone number and a fax number. Ideal support of user mobility is based on a unique identification of a user, which possibly remains unchanged for the user's lifetime.

Heterogeneous end systems. Future wireless networks will support a variety of end systems, which differ in their physical equipment. This equipment may include screens, video and sound support, input devices (touch screens, pointer, keyboards, voice recognition), data processing capabilities, storage and network devices. In particular it is expected that end systems will be equipped with more than one network interface in parallel. With the progress of software radio technology wireless interfaces are able to adapt to the current environment and switch between different modes if demanded/beneficial.

Heterogeneous access networks. Future wireless networks will be characterized by a variety of wireless access networks. It is expected that several wireless technologies will coexist. As a simple example, wireless LANs may offer high data rate service in indoor environments and 2nd/3rd generation cellular networks offer global coverage and low/medium data rate services. There is no wireless technology that provides all quality-of-service parameters all the time and there is a tradeoff between coverage, data rate and costs.

Protocol conversions. Although IP-based applications will dominate future traffic volume, existing standardized services (conventional voice, fax, old-style data applications) will still be supported by means of protocol conversions. This may include media conversions (e.g. fax-to-jpeg, email-to-voice, etc.) as well as well-known standardized supplementary services, such as GSM call forwarding, which have to be re-implemented in IP based applications. Moreover, not all application can be used in all end systems, e.g. due to limitations in equipment, data rate, etc. Therefore it is required to translate application-level protocols. The translation can be manyfold since they will be converted from a *preferred sender* type to a *preferred recipient* type.

Privacy. Two kinds of privacy is considered: Location privacy and anonymity. Since an IP address represents a host identifier as well as its physical location, the temporary IP address of a user betrays the users current location. Recent IP mobility approaches track the users reachability to maintain standard IP routing. But the distribution of location information could injure people's privacy. Future networks must offer mechanism to maintain location privacy to users who do not wish to reveal their current whereabout. Finally, the unique identification of a user could be used to compromise the user's privacy. Although the anonymity in the Internet is controversially discussed, it is one of the sources of the Internet's popularity.

Personal locator services. Future wireless networks will be provided with a locator service. This trend follows from the required provision of enhanced emergency services (e.g. E911 in the US), emerging locator technologies (e.g. GPS, assisted GPS). Moreover locator technologies are find growing acceptance and application (car navigation systems, active badges, etc.). Personal locator services in future wireless networks will determine and track the location of a user. The location information can be used to initiate a handover - in particular a vertical handover - by the network side. In the case of an upward vertical handover from a larger cell to a smaller cell classical handover initiation parameters fail and modified handover policies are required.

3.2 Conceptual consequences for handover design

The user requirements and technical opportunities of new technologies have a strong impact on the handover design. As a consequence we identify the following trends in support of handover:

Need for an enhanced address concept In current Internet networking, the IP address is used as a host identifier as well as for location information. To support mobility of IP end systems it is required to separate both meanings. Conceptually one solution is to hide the current location of the mobile from a correspondent host. Another might be to introduce a new address scheme into IP-based networks. The latter alternative is in line with the demand to include a user level into the addressing concept.

Support of vertical handover. Heterogeneous end systems and access networks lead to network structures that can be regarded as hierarchical cellular networks. Smaller cells are placed in hierarchical lower layer (see Fig. 3.1 for an example with a micro- and pico-cells). In such a network two categories of handover are possible: Handovers between cells of the same hierarchical level are called *horizontal* and handovers between cells at different hierarchical levels are called *vertical* handovers. A vertical handover is different from the horizontal case in some important respects. In particular a vertical handover is not symmetrical. An *upward* vertical handover is from a base station at a hierarchical lower layer to a higher layer (e.g. from a smaller cell to a larger cell). A *downward* vertical handover is from a base station at a higher layer to a lower layer. Since a higher layer offers larger spatial coverage than the lower layer, the upward vertical handover is more time critical: Suppose that a mobile moves out of a pico-cell cluster, then the handover duration to the micro-cell layer is time constrained. A downward vertical handover, on the other hand, is not necessarily time-critical, since connectivity to the old base station is available for the duration of the handover. Additionally for a vertical handover the quality of service can change. As an example the data rate can be decreased when changing from a pico-cell to a micro-cell layer or the mobile has to reduce its moving speed when changing to the pico-cell layer. Handover schemes have to satisfy the requirements of both handover types.

Differentiation of global and local mobility. Mobility can be considered as local and global mobility. The term *global* mobility (sometimes also referred to as *macro* mobility) describes the movement of end systems between geographical regions, whereas the term *local* mobility is used for the movement between neighboring base stations. For global mobility, maintaining connections (e.g. TCP connections) is usually not required. For local mobility, tearing down an ongoing connection is very undesirable (e.g. for rlogin sessions) and frequent handover might lead to considerable service degradation. The demand for a unified solution for global and local mobility is a misleading requirement and such an approach suffers from scalability problems.

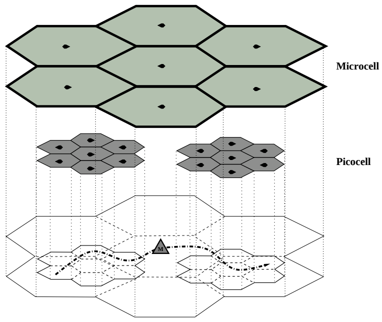

Figure 3.1: Hierarchical cellular network

Differentiation of active and idle end systems. Classical IP-based mobility approaches, such as IETF Mobile IP, do not differentiate between *active* end systems - which are currently sending or receiving - and *idle* end systems. Idle periods are very typical for traffic generated by IP-based applications (e.g. WWW sessions) and do not have timing constraints for service interruptions as active users [1] A handover scheme which is designed with respect to active users generate the same amount of signaling load for idle users. Thus, for a high number of users with a high percentage of idle users the handover scheme suffer from scalability problems.

Differentiation between state-full and state-free session. In general, Internet traffic can be divided in state-full and state-free- sessions, whereas the term *session* is regarded as an application-level concept. As an example, consider a state-free WWW session, which is composed of several parallel TCP connections. As the complement consider a *telnet* session with a long-lived TCP connection. For state-free sessions a rerouting of data flows (and possibly a forwarding of misdirected data) causes an unnecessary overhead. An automatic or manual restart of the session might be sufficient. For state-full sessions disruptions are usually very undesirable. A common solution for both cases causes overhead and a differentiation is expected to be more efficient.

[1] Note that an idle end system might have an open TCP connection.

Chapter 4

Overview of current handover approaches

In this chapter recent research-oriented handover approaches are described:

- IETF Mobile IPv4
- Extensions of IETF Mobile IPv4
- IETF Mobile IPv6
- Reverse Address Translation (RAT)
- Multicast-Based Handover
- HAWAII
- Cellular IP
- Mobile People Architecture
- ICEBERG
- Extended SIP Mobility.

In order to work out the basic assumptions behind the schemes, the description is organized according to the following structure:

- Motivation to develop a new approach
- Addressing concept
- Required mobility infrastructure
- Routing of packets
- Handover, in particular vertical handover
- Advantages and drawbacks of the approach

4.1 IETF Mobile IPv4

The IETF Mobile IPv4 [16, 17] is a well known approach for mobility support in IP networks and an accepted standard in the IETF community.

The motivation of Mobile IP is to offer a pure network layer solution for mobility support and to isolate higher layers from mobility. In particular, it aims at continuous TCP connections even though handover causes IP address changes. The IP routing mechanisms remain unchanged.

The fundamental assumption behind Mobile IP is that a mobile host owns an IP home address and gets assigned a temporary Care-of-Address (CoA) in a foreign network. A correspondent host addresses the mobile host via its IP home address. Mobile IP adds two new instances to the network infrastructure: A Home Agent (HA) and Foreign Agent (FA). Routing is performed by address translation and tunneling: Suppose a correspondent host (CH) wishes to send packets to the mobile host and sends it to its home address. The home agent intercepts and tunnels the packets to the CoA of the mobile host. To tunnel a packet it is encapsulated by the Home Agent. The Foreign Agent decapsulates the packets and forwards them via local mechanisms to the mobile host. For the reverse direction from the mobile to the CH, the mobile is allowed to send packets directly. This is referred to as *triangular routing*.

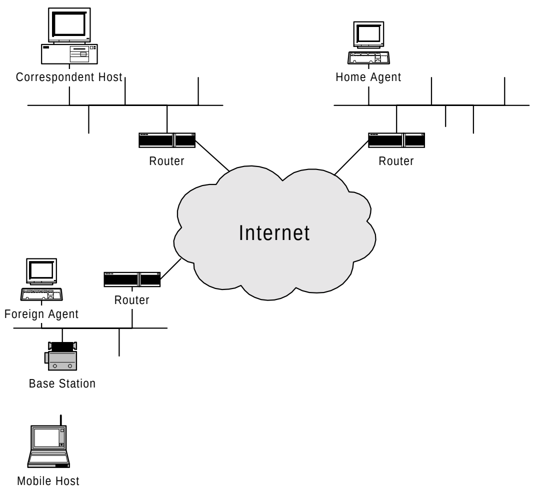

Figure 4.1: IETF Mobile IP network architecture

After the mobile detects that is has moved to a new IP subnet (e.g. by Lazy Cell Switching (LCS) [1], Prefix Matching [2], Eager Cell Switching (EGS) [3] [17]) and has obtained a new temporary IP address, it registers with its new IP Care-of-Address with the new Foreign Agent. The Foreign Agent in turn relays the registration to the Home Agent, which binds the new Care-of-Address to the mobile's IP home address. Following packets will be tunneled to the new Care-of-Address.

Considering a heterogeneous network architecture it is usually assumed that a cell cluster of the same layer forms an IP subnet with a single IETF Mobile IP Foreign Agent. Thus, for a handover

[1] The mobile detects the handover due to expiration of the last received agent advertisement

[2] The mobile receives a new advertisement with a different network prefix. This can be interpreted as a handover

[3] When the mobile receives multiple advertisements from different Foreign Agents it may select one of them.

between cells of different layers the Foreign Agent is changed and the same mechanism for vertical handover as for horizontal handover is applied.

When a mobile is registered with the Foreign Agent at the lower level (with smaller cells) a handover is detected by one of the common methods. The mobile registers with the new Foreign Agent and the Foreign Agent forwards the binding Update to the Home Agent (*Upward* vertical handover. For the *downward* vertical handover the mobile is registered with the Foreign Agent, which belongs to the cells at the higher layer. and the mobile moves into the cell coverage of the higher layer. In that case the Eager Cell Switching (ECS) method for initiating the handover does not work since the mobile continuously receives advertisements from the old base station. The other handover initiation algorithms (Prefix Matching, Lazy Cell Switching (LCS)) only work when the mobile has permanent connectivity to base stations of both layers. This is very inefficient. Nevertheless there is a problem: The mobile receives advertisements from Foreign Agents at both layers. This might lead to a frequent ping-pong handover. In summary, in the IETF Mobile IP approach no mechanism to initiate upward vertical handover exists and the mobile remains in the lowest hierarchical level, although there might be reasons to handover to the higher layer.

Although the IETF Mobile IP approach is very popular, it struggles with some serious problems:

- Each mobile end system is required to have an IP home address. In face of tight IPv4 address resources, this will lead to the usage of private IP addresses, which make the IETF Mobile IP scheme more complex.
- Ingress filters in routers drop IP packets that do not carry a topologically correct IP address. To avoid the loss of datagrams due to ingress filtering, reverse tunneling may be applied. The reverse tunneling in turn increases the routing overhead. The need for reverse tunneling is not automatically detected and dynamic switching between both tunneling modes is not IETF Mobile IP standard. For details see MosquitoNet's Extended Mobile IP approach [34].
- Indirect routing is considered as a serious drawback for traffic with stringent delay requirements, e.g. some RTP/UDP traffic, since it adds an additional end-to-end delay.
- Indirect routing adds data traffic to the network and tunneling increases the overhead.
- When IETF Mobile IP is widely used, the amount of signaling is large. Even for frequent local handovers, a binding update is generated each time the mobile changes the FA. When the Home Agent is distant from the mobile the generated signaling incorporates all network nodes along the route.
- For vertical handover, IETF Mobile IP applies the same scheme as for the horizontal case, which is independent of the hierarchical layer and for the upward and downward case. Therefore, a parameterization of IETF Mobile IP for optimal handover support (e.g. for fast handover) is difficult.
- IETF Mobile IP does not differentiate between active and inactive mobiles. In a scenario with IETF Mobile IP widely deployed, inactive user generate the same signaling traffic as active users, although it could be more efficient to locate them by searching (paging).

Partially these problems have been tackled by several Mobile IP extensions:

- Route optimization [18] solves the triangular routing problem by using *binding updates* to inform the correspondent host about the current temporary IP address. When a handover occurs the old Foreign Agent sends a *binding warning to the Home Agent which sends a* binding update to the Correspondent Host. The main drawback of this solution is that it requires a change in the IP protocol of the correspondent host since it must be able to encapsulate IP packets and to store Care-of-Addresses.

- The Mobile IP NAI extensions [4] can be used to uniquely identify the mobile. The NAI (Network Access Identifier) is usually provided by dial-in hosts which attempt to connect to a foreign domain with AAA servers (such as RADIUS [21] or DIAMETER [5]). The NAI is a unique identifier similar to the mobile permanent IP home address. First, the NAI can be regarded as an address supporting user mobility with Mobile IP. Second, when using the NAI the mobile's home address is *quasi* permanent.This allows to dynamically assign a Home Agent to the mobile. When the assigned Home Agent is in a Gateway router at the entry of the domain the triangular route of Mobile IP is identical with the optimal route. Nevertheless, the usage of an NAI results in a duplication of functionally.

4.2 Extensions of IETF Mobile IPv4

4.2.1 Hierarchical Foreign Agents

The approach with *Hierarchical Foreign Agents* [7, 8, 9] (http://www.cs.hut.fi/Research/Dynamics/) extends IETF Mobile IP. It addresses a drawback of Mobile IP: When the distance between the Foreign Agent and the Home Agent is large, the signaling delay for the registration may be long which results in long service disruption and packet losses.

In the approach with hierarchical foreign agents the addressing scheme is the same as in the traditional Mobile IP approach. Also this approach builds on the same blocks as the traditional Mobile IP. The difference is that the Foreign Agent functionality is distributed to several routers. These Foreign Agents can be configured in a tree-like structure. The *Highest Foreign Agent* is the root of the hierarchy, the *Lowest Foreign Agent* is close to the mobile on the path between the mobile and the Home Agent. *Intermediate Foreign Agents* are in between the path of the highest and the *Lowest Foreign Agent*. This Foreign Agent which belongs to the old and the new path is called the *Switching Foreign Agent*. The classification of the foreign agents is conceptual only. The hierarchy might collapse to a single foreign agent, as in the original IETF Mobile IP approach.

The approach works as follows: The *Lowest Foreign Agents* send announcements which include the own address and the address of the next higher level (and possibly all other levels including the *Highest Foreign Agent*). When a mobile first arrives at a visited domain, it sends a registration request to the *Lowest Foreign Agent* which creates an unacknowledged binding update and forwards the registration request upwards to the next higher Foreign Agent.

When a handover occurs the mobile generates a registration request which is forwarded by the *Lowest Foreign Agent*. At some point the *Switching Foreign Agent* receives the request and detects that a binding update already exists but is coming from a different *Lower Foreign Agent*. This is interpreted as a handover. The *Switching Foreign Agent* replies to the mobile with a *registration reply* message.

The main advantages of the approach with hierarchical foreign agents are:

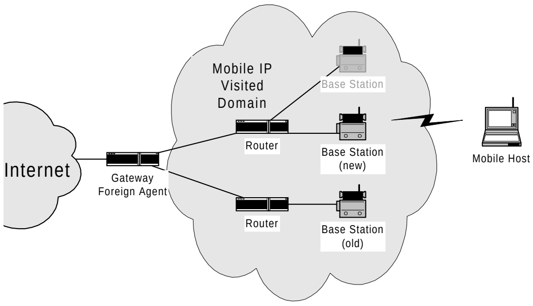

Figure 4.2: Network architecture for Mobile IP with hierarchical foreign agents

- The rerouting node is close to the mobile. This results in a shorter service disruption and less packet loss.
- The amount of signaling is reduced since less signaling data is sent to the Home Agent. This assumes that the lifetime of the binding is relatively long in order to avoid frequent binding refreshes sent from the mobile to the Home Agent.

A similar approach is investigated in the IETF [9] called *Regional Registrations*.

4.2.2 IETF Mobile IP Extensions by MosquitoNet

One of the extensions [33] addresses the issue of how IETF Mobile IP can be used most efficiently and flexibly on mobile hosts. IETF Mobile IP has been implemented with some modifications: Mobile IP specific routing table additionally to the regular IP routing table, extension of the Mobile IP Home Agent and of the protocol between mobile end system and Home Agent for registration. This allows to decide a) whether to use transparent mobility support or not and b) whether to use triangular routing or bi-directional routing. It is argued that Mobile IP implies some remarkable overhead which should be avoided when transparent mobility support by indirect routing is not necessary (a). Mobile IP Route Optimization (Triangular routing) fails when router ingress filtering is used: Packets are dropped, when they do not carry a topological correct IP source address. Therefore a mobile may use the more secure bi-directional tunneling, although it implies an additional overhead.

Essential in the context of this approach is the support of multiple interfaces in a mobile. Each interface carries a temporary IP address. A flow can be binded to a specific interface. This is done with the help of a socket option. For transmission the route lookup has been modified, so that only routes with that specific interface are considered. For reception of data, a flow-to-interface binding (flow is recognized by IP addresses and port number) is sent to the Mobile IP Home Agent, which forwards datagrams to the appropriate Care-of-Address. The Handover is similar to Mobile IP, but extends the protocol by an update of the Flow-to-Interface binding in the IETF Mobile IP Home Agent. The extension is mainly intended for vertical handover with mobiles with multiple parallel network interfaces.

4.3 IETF Mobile IPv6

In general the mobility support in IP version 6 [11] is based on the same main principles as Mobile IP for IP version 4. But in Mobile IPv6 a mobile is able to create its own Care-of-Address using its link-local address and automatic address configuration (combine advertised subnet prefix with own hardware address). Therefore there is no need for a foreign agent.

Mobile IPv6 introduces two new IPv6 Destination Options Header, namely a *Binding Update* and a *Binding Acknowledgment*. The destination options header is one of the so called *IPv6 extension headers* which is treated only by the final destination. The mobile can send directly a *Binding Update* in the same packets carrying effective traffic to its correspondent host (see Fig. 4.3). The correspondent host can then learn and cache the new mobile's Care-of-Address. As a result of this mechanism, when sending a packet to any IPv6 destination, a host must first check if it has a binding for this destination.

- If a cache entry is found, the host sends the packets directly to the Care-of-Address indicated in the binding, using an IPv6 Routing Header. This special extension header forces the datagram to follow a predetermined route which has two hops. The first hop is the Care-of-Address and the second hop is the home address of the mobile. This eliminates triangular routing. The mobile receives the packet and *forwards* it to the next hop specified in the routing header. The next (and final) hop is the home address of the mobile and the packet is *looped back* inside the mobile. Now the packet can be processed in the same way as if the mobile were at home.

- If no binding is found, the packet is sent to the mobile's home address. The Home Agent intercepts the packet and tunnels it to the Care-of-Aaddress as described previously.

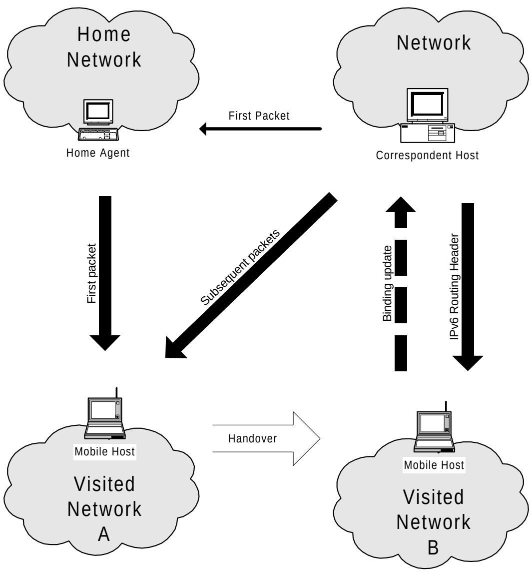

Figure 4.3: Binding update in Mobile IPv6

Compared to Mobile IPv4, some features have been integrated more efficiently int Mobile IP v6. The most relevant advantages are:

- For IPv6 there are enough IP addresses available. This eliminates the problem of address shortage in IP v4.
- Foreign Agents are not needed. In Mobile IPv6, mobiles utilize enhanced IPv6 features, such as address autoconfiguration [27] and neighbor discovery [15] for address configuration. DHCP is not required as well.
- Mechanisms for route optimization avoids triangular routing. While route optimitation is an additional functionality for Mobile IPv4, it is an integral part of Mobile IPv6.
- In Mobile IPv6 the problem of ingress filtering in routers (as known from Mobile IPv4) is solved: A mobile on a foreign link uses its Care-of-Address as source address of its packets, and includes its home address in the Home Address destination option. As the Care-of-Address is topologically correct, the packet will pass ingress filters.

4.4 Reverse Address Translation (RAT)

The RAT approach [24] (URL http://miphagent.mn.iscs.nus.edu.sg/rat/) is motivated by the small deployment of IETF Mobile IP. It is intended to simplify mobility support in order to break the *chicken and egg-trap* between the lack of applications, which require mobility support, and the poor deployment of IETF Mobile IP. The RAT approach can be considered as a tradeoff: On the one hand it dispenses with the requirement to maintain TCP connections. On the other hand overhead is decreased and most of the traffic can be routed directly. Moreover, the inventors of RAT argue that implementation of Mobile IP functionality is operating system dependent (e.g. registration, tunneling, etc.), whereas RAT aims at a solution, that is independent of the operating system.

In the RAT approach the mobile owns an IP home address and acquires a temporary IP address in the foreign network. The RAT approach adds new entities to the home network: a registration server and a RAT device. The network infrastructure remains unchanged. In particular there are no mobility specific entities in the foreign network required.

The RAT approach applies Network Address Translation (NAT) [25]. NAT is a Internet paradigm, which has been widely applied recently, e.g. for firewalls.

The RAT approach works as follows: Suppose a CH wishes to send a packet to the mobile host and directs it to the mobile's home address. In the home network, the RAT device intercepts the packet and performs a network address translation. Therefore it replaces the destination address with the mobile's temporary address and the source address with the address of the RAT device. Then the packet is send directly to the mobile without tunneling. In the reverse direction, the mobile sends a packet to the RAT device, which in turn performs the address translation and sends it to the correspondent host. This scheme is referred to as *reverse address translation*. One of the main advantages of this approach is that the indirect routing is deployed for CH initiated sessions only. When the mobile initiates the session, it will use its temporary address (which is topologically correct) and communicate with the CH directly [4] and no indirect routing via the home network is required.

The RAT approach does not address vertical handover specifically.

The RAT approach implies some drawbacks:

[4] The authors argue that most of the sessions are initiated by the mobile

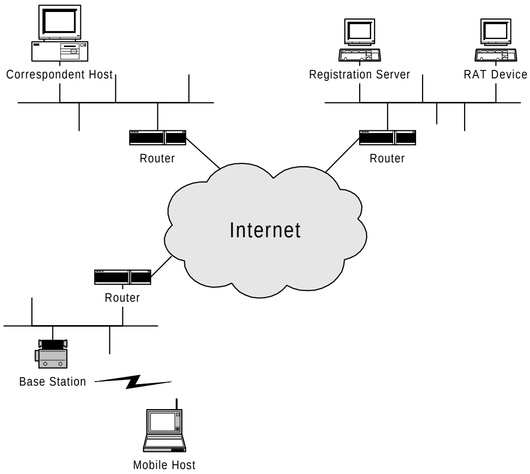

Figure 4.4: RAT network architecture

- No handover is possible by means of maintaining higher layer (e.g. TCP connections), since an address change leading to an interruption takes place. It is argued that most TCP connections for mobiles are short lived (e.g. WWW requests).
- ICMP payload translation is necessary, since they carry IP addresses.
- Recalculation of IP header checksum is necessary, since the IP address in the header is changed.
- Application Level Gateways (ALGs) are necessary, when applications use IP addresses or TCP/UDP ports in the payload (e.g. ping). Then this payload must also be changed.
- Some applications use IP address for authentication; with NAT this is not possible.

4.5 Multicast-Based Handover

In general IP multicast supports location-independent addressing and routing in IP networks. This ability is similar to the requirement of mobility support though in a different context. Thus the motivation of multicast-based handover is to re-use multicast mechanisms.

In the multicast-based handover approach a mobile gets assigned a temporary address, which is a unique IP multicast address. This address does not change for the lifetime of the session even when the mobile moves to a new IP subnet.

In multicast-based handover approach is at least one multicast router located in every IP subnet, where multicast services are offered. [5]. Multicast routers can be regarded as the mobility infrastructure, but the originally usage is efficient data distribution to a group of receivers.

The establishment of a session is based on multicast mechanisms and different from the unicast case: The mobile acquires a multicast address and joins the multicast group via registration at the

[5]It is not required that every IP subnet has got a multicast router, where multicast packets are delivered through: As an example DVMRP establishes tunnels between multicast routers, when an intermediate IP subnet does not has an multicast router

temporary multicast router. The multicast router in turn joins the multicast distribution tree which is constructed between the multicast routers with members of the particular multicast address with a multicast routing protocol (e.g. DVMRP, PIM-SM, etc.). The correspondent host sends packets with the mobile's temporary IP multicast address and the packets are distributed via the multicast distribution tree. In the reverse direction the mobile uses the (unicast) IP address of the correspondent host.

When a handover occurs, the mobile registers at the new multicast router with the same IP multicast address and the new multicast router joins the multicast distribution tree. The old multicast router leaves the multicast distribution tree (e.g. due to a time out or an explicite leave operation).

There are different sub-approaches to utilize multicast-based handover. In [14] it is intended to use today's IP multicast as it is available today in order to support handover. In [26] the IETF Mobile IP approach is extended by multicast: The Mobile IP Foreign Agents carry IP multicast addresses. When an handover event occurs packets are delivered efficiently from the Home Agent to at least two Foreign Agents and the handover latency can be decreased. In [6] an IP-style multicast is applied which realizes multipoint-to-multipoint communication in a switched access network. In this approach packets are distributed over an direct multipoint distribution tree of virtual circuits.

The advantages of the multicast-based approach are:

- The network node for the rerouting operation is located in this node, where the old and the new route diverge (and not e.g. in the home network distant from the mobile's temporary location).
- A dedicated mobile infrastructure in the network is not required, since the multicast infrastructure is re-used.
- It is not required that the mobile acquires a new multicast address when it has moved to a new subnet, whereas an IP multicast address is acquired only once for a session.
- For vertical handover a simultaneous distribution of packets to base stations at different hierarchical level decreases the handover latency. In that case the usage of multicast reduces the overhead for traffic distribution in the backbone network.

Nevertheless the multicast-based handover suffers from scaling problems, since today's IP multicast was designed to support efficiently *broadcast-like* applications, such as video distribution. However, the problems are mainly caused by the multicast itself and not due its utilization for handover. In particular the today's IP multicast

- has no indication of the receiver group size
- has no restriction of senders
- has high signaling overhead
- is complex

and therefore it is not generally available. Nevertheless it is commonly expected that IP multicast will be integrated more efficiently in the future Internet. Therefore it is expected that the IP multicast service model will change in order to support also *narrow cast* applications more efficiently which makes the deployment for handover more easy.

4.6 HAWAII

HAWAII [19, 20] stands for *H*andoff *Aw*are *W*ireless *A*ccess *I*nternet *I*nfrastructure. The HAWAII approach was proposed since Mobile IP results in high control overhead and high latency for local mobility. Also, in the case of a QoS enabled mobile host, acquiring a new Care-of-Address on every handover would trigger the establishment of a new resource reservation. The HAWAII approach extends Mobile IP and addresses these limitations.

HAWAII defines a *domain*. This is a division of the wireless access network under the administrative control of a single authority. The domain consists of routers and base stations. All of them are mobility-enabled by supporting HAWAII-specific signaling in order to optimize routing and forwarding. The router interconnecting the HAWAII domain and the Internet core network is called *Foreign Domain Root Router*. Each base station has Mobile IP Foreign Agent functionality. [6]

In the HAWAII approach mobility is separated between *intra-domain* handover and *inter-domain* handover and for both cases different mechanisms are defined. The first case is supported by HAWAII and the second case by Mobile IP.

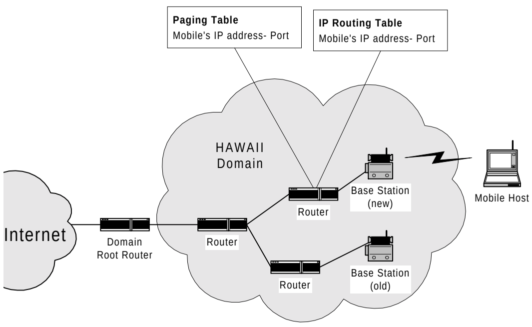

Figure 4.5: HAWAII network architecture

In the HAWAII approach a mobile has a home domain and a temporary unicast IP address. The home domain may support HAWAII. When the mobile is in a foreign HAWAII domain the temporary IP address is assigned once to the mobile and does not change as long as the mobile stays in the domain. Thus no address translation mechanisms is required and the Mobile IP Home Agent is not notified of the mobile's movement. Instead connectivity is maintained using dynamically established paths in the foreign HAWAII domain based on host entries in the routing table of selected routers. Thus, a HAWAII enabled access network does not rely on IP routing in the sense of routing based on the network's portion of the IP address. Instead the IP address is interpreted as a unique identifier and *not* as an location identifier.

As mentioned above, for global mobility support HAWAII reverts to traditional Mobile IP mechanisms. Let us first consider the case where the mobile is within the home domain. In this case the mobile carries a unicast IP address. [7] When the mobile powers up, it sends a Mobile IP *registration message* to the actual base station. The base station then propagates a HAWAII *path setup message* to the *domain root router* using a configured default route. Each router in the path between the mobile

[6] without decapsulation.
[7] The authors argue that this address might be quasi-permanent.

host and the domain root router adds a forwarding entry for the mobile. Finally, the domain root router acknowledges to the base station. The base station in turn replies the Mobile IP registration to the mobile. Packets for the mobile are sent to the domain root router based on the subnet's portion of the mobile's IP address. The packets are routed within the domain using the host-based forwarding entries. It is important to not that the entries are soft-state which are kept alive by periodic hop-by-hop messages.

When the mobile moves within the HAWAII domain the mobile registers with the new base station by sending a Mobile IP *registration request*. The new base station then sends a HAWAII *path setup update* message to the old base station. The old base station performs a routing table lookup for the new base station and adds a forwarding entry for the mobile's IP address. Then the message is sent to the upstream router. This router performs similar operations. When the router receiving this message is the crossover router [8] then this router adds a forwarding entry to the new base station and packets for the mobile are sent to the new base station. The path via the old base station will time out. This scheme is called *forwarding path setup* scheme since the HAWAII path setup update message is sent from the new to the old base station and the old base station forwards packets to the new base station for a limited time. This scheme is optimized for networks where the mobile listens/transmits to only one base station simultaneously. An alternative scheme is the the *Non-Forwarding* scheme, which is optimized for networks where the mobile is able to listen/transmit to two or more base stations simultaneously. In this path setup scheme the *path setup update* message travels from the new base station to the old base station via the crossover router. Thus packets are not forwarded from the old base station.

To interact with Mobile IP the mobile host is assigned a co-located Care-of-Address from its HAWAII foreign domain. A correspondent host directs the packets to the mobile's home address. The Mobile IP Home Agent intercepts the packets and tunnels them to the HAWAII *Foreign Domain Root Router* with the network portion of the outer IP address. This and the following router forwards the packets according to its host-based routing entries.

The HAWAII approach differentiate between active and idle users and appropriate states for the mobile host. For an active user the network knows the mobile's current base station and for an idle user the network knows only the base station approximately, such as a set of base station. When packets for an idle mobile arrives, the network "pages" the mobile to determine the mobile's current base station.

The main advantage of the HAWAII approach is that the rerouting node is the router where the path between the old and the new base station diverge and only these nodes are involved in processing a HAWAII path setup message. Additionally only the routers between the new base station and domain root router receive periodic refresh messages. Thus, HAWAII hides local mobility from the the Mobile IP Home Agent which results in less signaling traffic. Assuming a non-Mobile IP approach to support global mobility, for the HAWAII approach no tunneling is required.

[8] This router has a route to the old and the new base station via the same interface.

4.7 Cellular IP

The Cellular IP approach [28] (URL http://comet.columbia.edu/cellularip) envisions a networking environment with ubiquitous computers where highly mobile hosts often migrate during active data transfers and the users expect minimal disturbance to ongoing sessions. The authors argue that Mobile IP is no optimal solution, because it is optimized for macro-level mobility and relatively slowly moving hosts. Moreover, it is stressed that Mobile IP does not scale to a large number of mobile hosts, since every handover between Mobile IP Foreign Agents generates a binding update irrespective if the mobile is idle or active.

The Cellular IP approach proposes a hierarchical mobility management, which separates *global* from *local* mobility. For global mobility, IETF Mobile IP is applied to support handover across the Internet backbone. To support local mobility within the Cellular IP access network, regular IP routing is replaced with routing of packets hop-by-hop via lookup in specific tables. The tables apply soft-state principles, which are referred to as caches.

In a Cellular IP network a mobile is assigned a unique identifier, which is used to route packets. It is not required that a mobile has an IP Care-of-Address. For simplicity reasons the unique identifier is an IP address (e.g. home address) which makes inter-working with IETF Mobile IP more easy. However, this is not required, since within the Cellular IP access network no IP routing is performed.

For mobility support Cellular IP adds a *Gateway router* and *Cellular IP nodes* to the network infrastructure. A Gateway router interconnects the Internet backbone and the Cellular IP access network. The Cellular IP nodes are located in the Cellular IP access network and can be considered as base stations working at network level. It is not required that they are equipped with a wireless interface (if not they act as a regular network node).

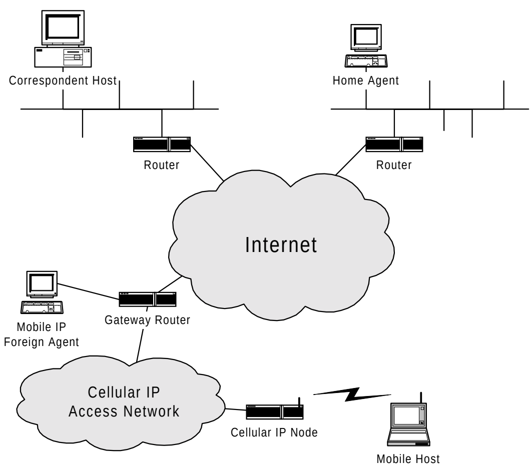

Figure 4.6: Cellular IP network architecture

The global mobility support in Cellular IP is provided straightforward by IETF Mobile IP. The Gateway router is co-located with a Mobile IP Home Agent and a Foreign Agent. The mobile registers the gateway's IP address with its Mobile IP Home Agent. Packets from a correspondent host are first routed to the Mobile IP Home Agent and then tunneled to the gateway. The gateway de-tunnels packets and forwards them towards the base stations. As long as the host is interconnected to the same access network, local mobility is hidden from the agent in the gateway router.

The local mobility support works as follows: Inside the Cellular IP access network, nodes are provided with a *Paging Cache* (PC) and a *Routing Cache* (RC). Both contain mappings between mobile host IDs and node ports (Output port similar to a router port) on a soft-state basis. Paging Caches are in a few nodes available. A Paging Cache is updated by data originating by the mobile (data packets or specific signaling packets). The paging cache is used to locate a mobile when there is no routing cache entry. In that case the Gateway Router caches the IP data packets in order to send a paging packet to the mobile across the Cellular IP nodes. The mobile replies to that paging packet and creates routing cache entries in every node along the route. Now, the cached IP packets can be sent along this route without address translation and tunneling. Paging Cache and Routing Cache entries are cleared by timers, with different timeout values: Routing Cache timeout is on the order of several IP packets, whereas the Paging Cache timeout is set according to the handover frequency. Thus, a idle and active mobiles can be managed separately with different data bases.

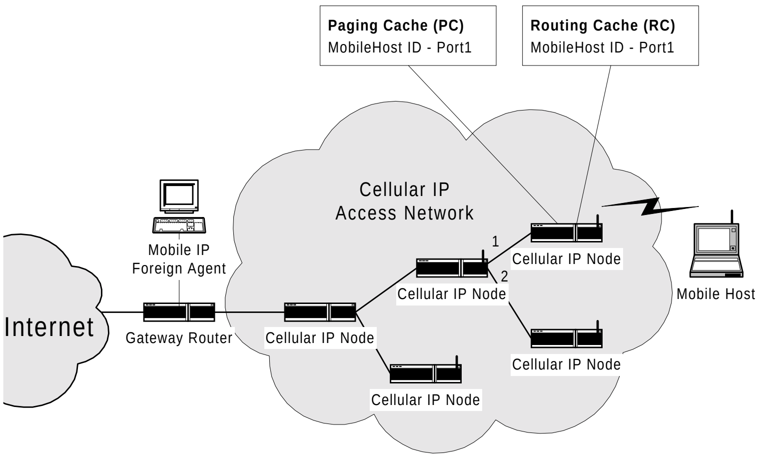

Figure 4.7: Cellular IP access network architecture

When a handover occurs two cases have to be considered. In the first case the mobile generates a *route update* packet when it enters the new cell in order to update the route caches in that nodes, where old and new route diverge. After the route caches are updated, data packets are sent to the new location of the mobile via the new route. For a limited time the old and the new routing cache entry can exist in the Routing Cache and data packets are sent via the old and the new route. This is used for *semisoft* handover. In the second case the Routing Cache entry in the Gateway was cleared triggered by a timer. Then a new paging packet is generated to locate the mobile. This explicite search causes a small delay in sending packets, but it allows longer timeouts decreases the amount of signaling packets.

For vertical handover two cases have to be considered: a) When both wireless cells do not belong to the same Cellular IP access network the handover is just a regular Mobile IP handover and b) When both wireless cells belong to the same the same Cellular IP access network the handover is handled by the Cellular IP handover scheme with its advantages to IETF Mobile IP.

The Cellular IP approach offers several advantages:

- Cellular IP nodes are relatively simple and configure itself.

- It differentiates between active and idle mobiles. For active mobiles the Routing Cache is used, whereas for idle mobiles the Paging Cache is applied. Idle mobiles generate less signaling traffic, since they update the cache of the nodes less frequently than active mobiles.
- When an idle mobile moves to a new Cellular IP node only the caches of these nodes are updated where the old and the new path diverge. Therefore a joint strategy of registration and searching is applied. This keeps the network free of signaling load.
- The soft state of caches with timeout values enables optimal tuning to minimize signaling load which improves handover performance. Additionally it increases error tolerance.
- Very flexible: It scales from a small office environment to a wireless access network. Due to the separation of idle and active mobiles the administration of a high number of mobiles is possible.
- Mobile-to-mobile communication is routed via the gateway (and not via the Mobile IP Home Agent).

Nevertheless the Cellular IP approach assumes that every node is mobility aware, since it replaces regular IP routing with *Routing Cache* lookups and signaling to locate mobiles using *Paging caches*. (Note that when a nodes is not provided with a RC and PC, they broadcast data to all ports.)

4.8 The Mobile People Architecture

The main goal of the Mobile People Architecture [1, 13, 22] (URL http://mosquitonet.stanford.edu) is to maintain person-to-person reachability while preserving the mobile's person privacy.

In the Mobile People Architecture a user is identified by a *Personal Online ID*. Additionally, a user is addressed by *Application Specific Addresses*. Mobility is supported by mapping the Personal Online ID to Application Specific Addresses.

In the Mobile People Architecture a new entity is added to the network: It is called a *Personal Proxy* and acts as a *person level router*. (The person level is added to the communication layer model on top of the application level.) The Personal Proxy tracks the user's current reachability, converts media and forwards data to a specific end system. It is located in the mobile's home network (if any) or is offered by a trusted third party server.

When a user wishes to communicate with the mobile *person* a call (Call is regarded as a kind of session) is directed to the Personal Proxy and then to the mobile person's preferred end system. When the reachability of the mobile person changes, the proxy state is updated by the tracking agent. The update can be done in a scheduled manner, manually or automatically.

It is assumed that local mobility is handled within the access network and hidden from the Personal Proxy. The case that a user changes its end system can be regarded as vertical handover. Then the user updates its Personal Proxy (manually or automatically) and new calls will be directed to the user's new application specific address (ASA). The case, that a user changes the ASA while receiving service is not considered.

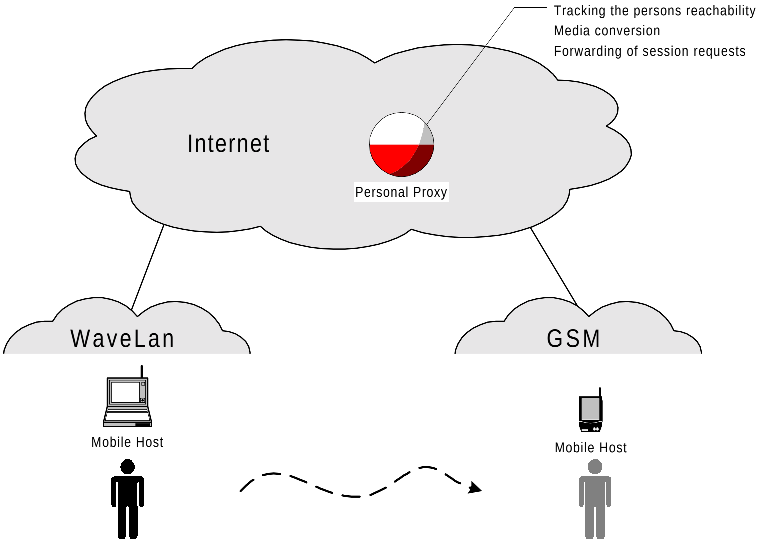

Figure 4.8: The Mobile People network architecture

4.9 ICEBERG

The motivation of the ICEBERG project (**I**nternet **C**ore **BE**yond the thi**R**d **G**eneration) [29] (URL http://iceberg.cs.berkeley.edu) is the current diversity of access networks, end systems, and services; in particular, traditional telephony services and data services. Therefore the ICEBERG project aims to support personal mobility in the sense of seamless access to services, independent of the access network and end system. It is intended to give the control of a communication to the callee, and not to the caller.

In ICEBERG, a user can be uniquely identified (*unique-id*). Additionally, the user is associated with one or several service IDs (e.g. phone number, email address, IP address). To achieve mobility a *unique-id* is mapped to the *service-id*.

In general, the ICEBERG network architecture consists of the Internet Core and several different access networks (e.g. GSM, PSTN, WLAN). At the interface between the core network and an access network an IAP (Iceberg Access Point) transforms services (media converter). Additionally ICEBERG adds service agents to the core network: preference registries, personal activity tracker (PAT) and extended naming services. The preference registry stores user preference profiles, which can be modified by user interaction or by the personal activity tracker (PAT), which gives inputs about location information.

Suppose a correspondent user wishes to call the mobile user. The call is routed to the IAP. In the access point a name service lookup is performed, the preference registry of the called user is located and the preferred end system is determined. After then the call is established via the correspondent interface. A service conversion (e.g. fax to jpeg) is executed in the IAP.

The ICEBERG approach focuses on user mobility between several access networks. It is implicitly assumed that host mobility is supported transparently in the access networks by technology specific handover schemes (e.g. for GSM, IEEE802.11, etc).

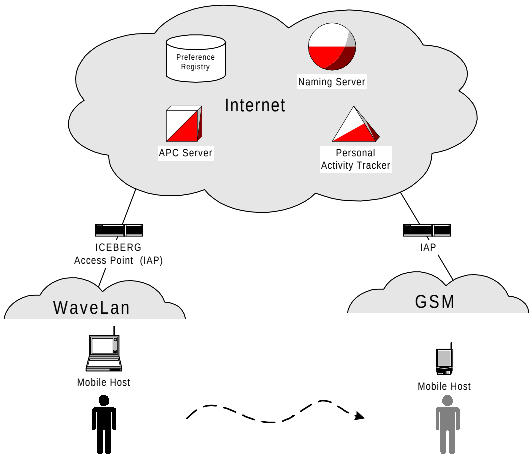

Figure 4.9: ICEBERG network architecture

4.10 Extended SIP Mobility

SIP [10] stands for ***S**ession **I**nvitation **P**rotocol* and *SIP Mobility* [30] is an mobility approach, which utilizes application level signaling. The motivation of the *Extended SIP mobility* can be found in drawbacks of the IETF Mobile IPv4 approach. The authors argue that for real-time traffic over IP , which is mostly RTP [23] over UDP traffic there is a need for fast handover, low latency and high bandwidth utilization. IETF Mobile IPv4 suffers from indirect communication which increases the delay and causes an overhead due to tunneling, which decreases the bandwidth utilization.

The *Extended SIP mobility* approach introduces mobility awareness at a higher layer than the network layer. SIP already supports user mobility and the approach is to extend SIP as an application layer signaling protocol in order to support *end system mobility*.

The main assumption behind the *Extended SIP mobility* approach is, that a mobile user is identified by a unique address (e.g. user@realm). This unique address is mapped to the current IP address of the mobile user's end system. No explicite home IP address is required. SIP introduces a SIP agent at the user's side and a SIP server (SIP redirect server or SIP proxy server) and location server to the network infrastructure.

The *user* mobility is supported as follows: Suppose a user wishes to initiate a session, an invitation is directed to the SIP server, which in turn queries the location server for the current IP address of the mobile user's end system. The SIP server sends the invitation to the called user. The invitation contains the IP address of the callee. If the mobile user moves, the location server is updated and new sessions will be set up to that new IP address.

End system mobility to that scheme is mainly understood as an increased roaming frequency and as a change of an IP address during an ongoing session. Assuming that a session is already established, then the mobile registers with the location server the new temporary address and the mobile re-invites the correspondent host with the same session identifier and the new temporary address (in the *Contact field of the SIP message*. The session can be continued, although the IP address has changed.

It is important to note that SIP does not support TCP. Therefore *Extended SIP mobility* supports UDP traffic only. For TCP traffic it is proposed to use Mobile IP. It is argued that both approach can

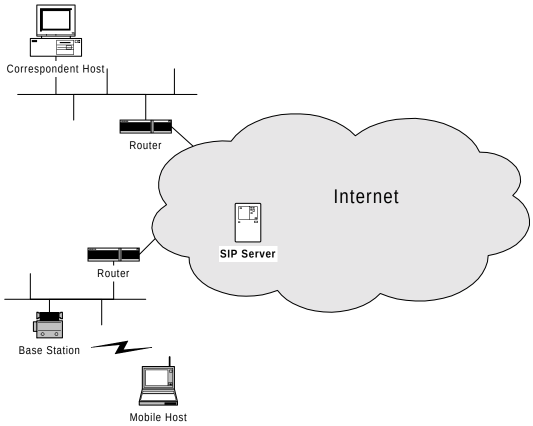

Figure 4.10: Network architecture in the *Extended SIP mobility* approach

coexist: for TCP traffic Mobile IP is applied and for UDP traffic the Extended SIP mobility approach. For the simultaneous usage the MosquitoNet approach of a *Mobile Routing Table* [34] is adopted.

Chapter 5

Summary and Conclusions

This chapter compares the handover approaches described in the last chapter. Foremost the approaches are compared with respect to the general mobility problem. Then the mechanisms are summarized which are used by the particular handover approaches. For the comparison it is assumed that handover can be generalized as a composition of several mechanisms.

As stated previously the general mobility problem can be regarded as an addressing and routing problem. Generally spoken, the problem is solved by breaking the double meaning of the IP address as an endpoint identifier and a location identifier. Three main different schemes with respect to the address assignment have been identified:

1. Permanent unicast IP address and temporary unicast IP address
2. Multicast IP address
3. Permanent non-IP address and temporary unicast IP address

In the first addressing scheme the mobile carries two IP addresses - a unicast IP address which identifies the mobile uniquely in the network and a unicast IP address which identifies the current location. The home address is permanent, whereas the foreign address changes. The addresses are assigned to each other by a *address translation mechanism*. In IETF Mobile IP (both for IP version 4 and version 6), the address assignment is provided by IP-in-IP tunneling and the address change is transparent to higher protocol layers. In the RAT approach, network address translation is applied and the address change is visible to higher protocol layers.

In the second addressing scheme, the mobile carries a multicast IP address. This address identifies the mobile and *not* its current location. The IP multicast address remains unchanged since it does not reflect the location. Hence *no* address translation mechanism as in the first case is required and higher layers are not effected by any address change.

In the third scheme the user is identified by a unique identifier which is usually a non-IP address (e.g. user@realm). The mobile carries a temporary unicast IP address. The unique identifier is permanent, whereas the temporary unicast IP address changes. The unique identifier is used to identify the user uniquely in the network, whereas the temporary unicast IP address identifies the location. An address translation mechanism is required to map the unique identifier to the temporary IP address. Four approaches apply this scheme: *Mobile People Architecture*, *ICEBERG*, *Cellular IP*, and *Extended SIP Mobility*. The first and the second approach assume a very general architecture which include also non-IP access networks. Nevertheless their approach is valid for *All-IP* networks when the temporary addresses are restricted to IP addresses. In the *Mobile People Architecture* the unique identifier is called *Personal Online ID*. This ID is assigned to an *Application Specific Address* of the *Personal Proxy* (for privacy reasons) or the direct address of the user. The *Application Specific Address* is ultimately an IP address [1]. For address mapping the application (or the user) query a *directory service*. In the *ICEBERG* approach the user is identified by a *unique ID*. This ID is mapped to a *service-specific-id* using a *naming service* or the *user's preference registry*. The address translation (in the ICEBERG context *Naming Service* is again a DNS-like mechanism which unifies several addressing concepts. In *Cellular IP* a unique identifier is assigned to the mobile which is used for routing purposes. This identifier is quasi permanent in the Cellular IP access network. If Cellular IP is used for local mobility and Mobile IP for global mobility then the Mobile IP address is permanent and the temporary IP address is used as a unique identifier in the Cellular IP access network. In the *SIP mobility* approach a user is identified by a unique address which is mapped to a temporary IP address. Both addresses are mapped via a SIP server.

Also, in the HAWAII approach, the unicast IP address is used like a unique identifier. The mobile carries a unicast IP address, therefore it can be assigned to the first category. But this address does not reflect the mobile's location since the routing operation is not base on the network portion of the IP address. Therefore, it can also be subsumed into the last category.

Routing can be regarded as an orthogonal component to addressing: Either the addressing allows

1. Direct routing
2. Indirect routing
3. Hybrid (direct and indirect) routing

Direct routing provides an optimal route between a correspondent host and a mobile. In principle direct routing can be provided with multicast addresses and in schemes utilizing a unique identifier for locating. Nevertheless this statement is true only in principle: For the multicast addressing scheme most multicast distribution trees provide non-optimal routing (such as core-based trees), which has mainly scalability reasons. Also, the address scheme which maps the unique identifier to the temporary IP address does not always provide direct routing since these approaches place a proxy server into the network in order to ensure privacy and/or perform application-specific routing and packet processing. In the *Cellular IP* approach packets are routed with the unique identifier, hence no IP routing is applied. As a consequence, Cellular IP can only be applied in Cellular IP access networks.

Indirect routing is - in principle - provided by *IETF Mobile IP* and Mobile IP with hierarchical foreign agents. Again, for this statement some additional remarks are required: Packets from the mobile to a correspondent host can be sent directly (triangular routing). But as already mentioned, this solution suffers from packet drops due to ingress filtering. Likewise, *Mobile IP with hierarchical foreign agents* does not modify the routing of packets via the Mobile IP Home Agent.

Hybrid routing provides both routing schemes - direct as well as indirect routing. The *RAT* approach supports direct routing for mobile initiated sessions and the indirect routing for correspondent host initiated sessions. The MosquitoNet extensions of IETF Mobile IP allows the mobile end system to select between transparent routing (Mobile IP style) and direct routing. IETF Mobile IPv6 bases

[1] More precisely, an intermediate step is required to map the *Application Specific Address* to the IP address with a standard service (e.g. Domain Name Service DNS)

in general on the indirect routing of Mobile IPv4, but integrates route optimization in order to avoid triangular routing.

Table 5.1 summarizes how the particular approaches solve the general mobility problem.

Addressing vs. Routing	**Unicast IP Addresses**	**Multicast IP Address**	**Non-IP Address**
Direct Routing	HAWAII	Multicast-based approach	Mobile People Architecture ICEBERG Cellular IP Extended SIP mobility
Hybrid Routing	RAT MosquitoNet Extended Mobile IP IETF Mobile IPv6		
Indirect Routing	IETF Mobile IPv4 MIP with hierarchical foreign agents		

Table 5.1: Comparison of handover approaches with respect to the general mobility problem

Table 5.2 summarizes handover-related functions of the approaches which where described in chapter 4. The first four rows compare the mechanisms (*Detection of link availability, registration, registration update and the address translation mechanism*). The item *Database for location information* mainly indicates if the approach employs a centralized or distributed database. The item *Rerouting node* describes if the node where the rerouting takes place is optimal or not. *Support of user mobility* denotes if a user mobility concept is already included in the approach. *Support of multiple interfaces in the mobile* and *Simulcasting to multiple base stations* refers to the support of vertical handover. The items *Differentiation between active and idle users* and *Differentiation between state-full and state-less session* compares if these mechanisms are used in the particular approaches. Finally, *Location privacy* shows if a correspondent host is able to determine the location of a mobile revealing the user's privacy.

We envision a future generation wireless network as a pure IP-based network. Base stations and routers in the access network are IP addressable entities. We believe that mobility should be handled at network layer in cooperation with link layer and application layer.

The *IETF Mobile IP* framework is a strong candidate for applications in future *All-IP* wireless networks. Several drawbacks of the basic Mobile IP were listed in Section 4.1. Some of them are tackled by extensions of the basic approach, such as improved support of local mobility, support of user mobility by NAI, route optimization, and others.

But we believe that a unified mechanism for all mobility scenarios is a misleading requirement. Although we are convinced that Mobile IP will find its deployment for global mobility support we favor a concept which employs direct communication between hosts as far as location privacy is not violated. Therefore, a new addressing concept is required in any case. The usage of Network Access Identifiers in Mobile IP is one possible concept since NAIs are popular for dial-in users. Nevertheless the assignment of a NAI *and* a permanent IP address which in turn is mapped to a temporary address lead to a three-tier addressing scheme where one level might be redundant. Moreover, as characterized in this report, many of the today's Internet applications are based on a client-server paradigm which do not require mobility support. Putting it briefly, for global mobility we advocate a addressing scheme which maps user-level addresses to temporary IP addresses.

It is commonly expected that some future applications (audio and video) will require handover support. In that case a mobile may use a mobility scheme supporting local mobility in a domain with heterogeneous wireless technologies. Nevertheless, it can be estimated that a number of open questions regarding handover policies for vertical handover need to be answered.

Finally, *Multicast-based handover* offers a number of advantages for handover support: No handover-specific signaling is required, the rerouting node is close to the base station and mechanisms to optimize handover can be realized easily. Today's multicast has some drawbacks and it is not generally available. Nevertheless, it is commonly expected that an optimized multicast will be integrated in the future Internet, such as *small group multicast*. Nonetheless, a number of open questions exists and some extensions to the basic multicast can be useful.

Table 5.2: Comparison of handover approaches with respect to general functions

General functions	Basic Mobile IPv4 *(v6)*	Mobile IP with Hierarchical FA	MosquitoNet's Extended Mobile IP	Multicast-based handover	RAT	Cellular IP	HAWAII	Mobile People Architecture	ICEBERG	Extended SIP Mobility
Detection of new link availability	FA advertisement / solicitation *(router)*	Similar to basic Mobile IP	Similar to basic Mobile IP	IGMP advertisement by multicast router	Access network specific	Access network specific	Access network specific	Dependent on policy	Dependent on policy	Access network specific
Registration	At FA and HA *(At HA)*	At FA(s) and HA	At HA (Co-located FA)	Multicast join operation	At registration server in home network	Once at HA, route update for active hosts	Once at HA, path setup for active hosts	At MPA's personal proxy	At preference registry	At SIP Server
Registration update	Registration and binding update at HA *(and CHs from BU list)*	Regional registration at FA(s)	Similar to basic Mobile IP	Multicast join-/leave-operation	At registration server in home network	Route update towards the gateway router	Path setup update for active hosts towards Domain Root Router	At MPA's personal proxy	At preference registry	At SIP Server
Database for location information	Registration table in Home Agent *(and binding cache in CHs)*	Registration table in HA and tables in FAs	Similar to basic Mobile IP	Distributed in multicast routers	Registration server	Routing and paging caches	Host-based routing table entries and paging caches	Personal Proxy	Preference registry, naming server	SIP server
Address translation	Encapsulation	Encapsulation	Similar to Mobile IP or direct communication	None	NAT	None	None	Directory service	Directory service	Via SIP server
Rerouting node	HA in home network *(or CH directly)*	Switching FA in visited domain	Similar to basic Mobile IP	Multicast router close the base station	Registration server in home network	*Inter domain*: Home Agent *Intra domain*: Node close to mobile	*Inter domain*: Home Agent *Intra domain*: Router close to mobile	MPA's personal proxy	ICEBERG Access Point (IAP)	CH / Mobile
Support of user mobility	NAI extensions	Similar to basic Mobile IP	Similar to basic Mobile IP	No	No	No	No	Yes	Yes	Yes
Support of multiple interfaces in mobile	Yes, with multiple IP addresses	Similar to basic Mobile IP	Yes, with multiple IP addresses	Yes, with same IP multicast addresses	No	NA	NA	Yes	Yes	Yes
Simulcasting to multiple base stations	Yes, with Simultaneous Binding Option	Similar to basic Mobile IP	Yes, simultaneous binding and flow-to-interface-binding in additional to basic Mobile IP	Yes	No	Yes	No	No	No	No
Differentiation of active and idle hosts	No	No	No	No	No	Yes, paging for idle hosts	Yes, paging for idle hosts	No	No	No
Differentiation between state-full and state-free sessions	No	No	No	No	No	No	No	No	No	No
Location privacy	Yes	Yes	Selectable	Yes	Mobile initiated sessions: No CH initiated sessions: Yes	Yes	Yes	Yes	No	Yes

Bibliography

[1] G. Appenzeller, K. Lai, P. Maniatis, M. Roussopoulos, E. Swierk, X. Zhao, and M. Baker. The Mobile People Architecture. Technical report, Stanford University, 1996.

[2] A. Bakre and B. Badrinath. Implementation and Performance Evaluation of Indirect TCP. *IEEE Transaction on Computers*, 46(3):279–289, 1997. `http://computer.org/tc/tc1997/t3toc.htm`.

[3] H. Balakrishnan, S. Seshan, and R.H. Katz. Improving Reliable Transport and Handoff Performance in Cellular Wireless Networks. *ACM Wireless Networks*, 1(4):469–481, 1995. `http://daedalus.cs.berkeley.edu/publications/winet.ps.gz`.

[4] P. Calhoun and C. Perkins. Mobile IP Network Access Identifier Extension for IPv4. RFC 2794, March 2000. `http://www.ietf.org/rfc/rfc2794.txt`.

[5] P. Calhoun, A. Rubens, H. Akhtar, and E. Guttman. DIAMETER Base Protocol. INTERNET DRAFT work in progress, December 1999. `http://www.ietf.org/internet-drafts/draft-calhoun-diameter-12.txt`.

[6] A. Festag, L. Westerhoff, A. Assimakopoulos, and A. Wolisz. Rerouting for Handover in Mobile Networks with Connection-Oriented Backbones - An Experimental Testbed. In *Proc. of ICATM'2000*, pages 491–499, June 2000. `http:/www-tkn.ee.tu-berlin.de`.

[7] D. Forsberg, J.T. Malinen, J.K. Malinen, and H.H. Kari. Increasing communication availability with signal-based mobile controlled handoffs. In *Proc. of IP based Cellular Networks (IPCN2000)*, Paris, France, May 2000. `http://www.cs.hut.fi/Research/Dynamics/publications/dynamics_ipcn2000_p%aper.ps`.

[8] D. Forsberg, J.T. Malinen, T. Weckstroem, and M. Tiusanen. Distributing Mobility Agents Hierarchically under Frequent Location Update. In *Proc. of Sixth IEEE International Workshop on Mobile Multimedia Communications (MOMUC'99)*, San Diego, CA, USA, 1999. `http://www.cs.hut.fi/Research/Dynamics/publications/dynamics_momuc99_pa%per.ps`.

[9] E. Gustavsson, A. Jonsson, and E.C. Perkins. Mobile IP Regional Registration. INTERNET DRAFT work in progress, July 2000. `http://www.ietf.org/internet-drafts/draft-ietf-mobileip-reg-tunnel-03.t%xt`.

[10] M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberger. SIP: session initiation protocol. IETF Request for Comments (Proposed Standard) 2543, May 1999.

[11] D. Johnson and C. Perkins. Mobility Support in IPv6. INTERNET DRAFT work in progress, April 2000. `http://www.ietf.org/internet-drafts/draft-ietf-mobileip-aaa-key-01.txt`.

[12] D.A. Maltz and P. Bhagwat. MSOCKS: An Architecture for Transport Layer Mobility. In *Proc. of IEEE INFOCOM'98*, pages 1037–1045, San Francisco, CA, USA, March/April 1998. `ftp://ftp.monarch.cs.cmu.edu/pub/dmaltz/msocks-infocom98.ps.gz`.

[13] P. Maniatis, M. Roussopoulos, E. Swierk, M. Lai, G. Appenzeller, X. Zhao, and M. Baker. The Mobile People Architecture. *ACM Mobile Computing and Communications Review (MC2R)*, July 1999. `http://mosquitonet.stanford.edu/publications.html`.

[14] J. Mysore and V. Bharghavan. A New Multicast-based Architecture for Internet Mobility. In *ACM MOBICOM 97*, October 1997. `http://www.acm.org/pubs/articles/proceedings/comm/262116/p161-mysore/p1%61-mysore.pdf`.

[15] T. Narten, E. Nordmark, and A. Simpson. Neighbor Discovery for IP version 6 (IPv6). RFC 2461, December 1998. `http://www.ietf.org/rfc/rfc2461.txt`.

[16] C. Perkins. IPv4 Mobility Support. Request for Comments 2002, October 1996.

[17] C. Perkins. *Mobile IP Design Principles and Practices*. Number ISBN: 0-201-63469. Addison-Wesley Longman, Reading, MA, USA, 1998.

[18] E.C. Perkins and D.B. Johnson. Route Optimization in Mobile IP. INTERNET DRAFT work in progress, February 2000. `http://www.ietf.org/internet-drafts/draft-ietf-mobileip-optim-09.txt`.

[19] R. Ramjee and T. LaPorta. Paging support for IP mobility. INTERNET DRAFT work in progess, July 2000. http://www.ietf.org/internet-drafts/draft-ietf-mobileip-paging-hawaii-01.txt.

[20] R. Ramjee, T. LaPorta, S. Thuel, and K. Varadhan. IP micro-mobility support using HAWAII. INTERNET DRAFT work in progess, July 2000. `http://www.ietf.org/internet-drafts/draft-ietf-mobileip-hawaii-01.txt`.

[21] C. Rigney, A. Rubens, W. Simpson, and S. Willens. Remote Authentication Dial In User Service (RADIUS). Request for Comments (Proposed Standard) 2138, April 1997. `http://www.ietf.org/rfc/rfc2138.txt`.

[22] M. Roussopoulos, P. Maniatis, E. Swierk, K. Lai, G. Appenzeller, and M. Baker. Person-Level Routing in the Mobile People Architecture. In *Proceedings of the USENIX Symposium on Internet Technologies and Systems*, October 1999. `http://mosquitonet.stanford.edu/publications.html`.

[23] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RTP: a transport protocol for real-time applications. IETF Request for Comments (Proposed Standard), January 1996.

[24] R. Singh, Y.C. Tay, W.T. Teo, and S.W. Yeow. RAT: A Quick (And Dirty?) Push for Mobility Support. In *Proceedings of the Second IEEE Workshop on Mobile Computing Systems and Applications*, New Orleans, LA, USA, Febrary 1999. `http://www.research.att.com/conf/wmcsa99/papers/singh.ps.gz`.

[25] P. Srisuresh and K. Egevang. Traditional IP Network Address Translator (Traditional NAT). INTERNET DRAFT draft-ietf-nat-traditional-04.txt, April 2000. `http://www.ietf.org/html.charters/nat-charter.html`.

[26] M. Stemm and R. Katz. Vertical Handoffs in Wireless Overlay Networks. *ACM Mobile Networking (MONET), Special Issue on Mobile Networking in the Internet*, Winter 1998. `http://HTTP.CS.Berkeley.EDU/~stemm/publications/monet97.ps.gz`.

[27] S. Thomson and T. Narten. IPv6 stateless address autoconfiguration. INTERNET RFC, December 1998. `http://www.ietf.org/rfc/rfc2462.txt`.

[28] A. Valko. Cellular IP - A New Approach to Internet Host Mobility. *ACM Computer Communication Review*, January 1999 (to appear).

[29] H.J. Wang, B. Raman, C. Chuah, R. Biswas, R. Gummadi, B. Hohlt, X. Hong, E. Kiciman, Z. Mao, J.S. Shih, L. Subramanian, B.Y. Zhao, A.D. Joseph, and R.H. Katz. ICEBERG: An Internet-core Network Architecture for Integrated Communications. *IEEE Personal Communications*, (Special Issue on IP-based Mobile Telecommunication Networks.), 2000. `http://iceberg.cs.berkeley.edu/publications.html`.

[30] E. Wedlund and H. Schulzrinne. Mobility Support using SIP. In *Proc. of Second ACM/IEEE International Conference on Wireless and Mobile Multimedia WoWMoM99*, Seattle Washington, USA, August 1999.

[31] M. Weiser. The Computer for the 21st Century. *Scientific American*, 3(265):94–104, 1991. http://www.ubiq.com/hypertext/weiser/UbiHome.html.

[32] A. Wolisz. Wireless Internet Architecture: Selected Issues. In *Proc. of Personal Wireless Communications*, Gdansk, Poland, September 2000. `http://www-tkn.ee.tu-berlin.de/publications/proc.html`.

[33] X. Zhao, C. Castelluccia, and M. Baker. Flexible Network Support for Mobile Hosts. In *To appear in MONET Special Issue on Management of Mobility in Distributed Systems*, 1999. `http://mosquitonet.stanford.edu/publications.html`.

[34] X. Zhao, C. Castelluccia, and M. Baker. Flexible Network Support for Mobility. In *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom 1998)*, Dallas, USA, October October 1998. `http://mosquitonet.stanford.edu/publications/flexible.ps`.