Thesis proposal: Supporting real-time services to mobile Internet hosts

Jon-Olov Vatn
Telecommunications Systems Lab (TSLab)
Department of Microelectronics and Information Technology (IMIT)
Royal Institute of Technology

1.0 Background, motivation and statement of the proposed problem

1.1 Trends and vision

With the introduction of IP based telephony services, the Internet has started to challenge the traditional PSTN networks as an infrastructure for providing real-time interactive services. This upcoming paradigm shift is not only driven by the desire to provide cost efficient solutions, but by basing the communication on IP we expect that the end-users will experience a greater set of attractive services over a single connection compared to what is provided by a PSTN today.

Looking a little further ahead, mobile communication systems will also become IP based. Companies, universities and private persons have started to extend their local area networks to provide wireless access by attaching wireless access points (APs) to their LAN. Wireless ISPs (WISPs) are putting up wireless LAN (WLAN) APs at public hot spots, thereby providing a complement or even a competitive alternative to the wireless WANs (WWANs) being developed and deployed today. As more and more people start to communicate using WLAN access, they will naturally wish to use this infrastructure for interactive real-time applications, such as mobile telephony. Thus, we can expect that there is a need to provide terminal mobility, even in a heterogeneous access network environment, where the heterogeneity both in terms of operators and wireless technologies, while IP (v4 or v6) constitutes the common platform.

But with the desire to use the Internet as a basis for mobile telephony comes the challenge to provide this in a way that meets the expectations from the users as well as the providers. Just as with ordinary IP telephony, the problem of providing an acceptable level of quality in a packet based network has to be faced. But in addition to this there is a set of issues that are specific to the use of wireless access and mobility support. The user (or more specifically, the user's terminal) should be able to make a handover between different APs without losing ongoing communication sessions. These handovers should be possible not only within a single operator's network, but also between different operators based on roaming agreements. To provide coverage even outside the urban hot-spots, handovers could also occur between WLAN and WWAN networks. While it is important that these handovers are performed without unacceptable disruption of ongoing communication sessions, it is also important for both the users and the operators that access control is performed in a secure manner to enable support for services such as accounting and to avoid misuse of the system. Other properties of wireless and mobile networks that add to the difficulty of providing real-time services are, e.g., the non-optimal routing commonly present in these systems, and higher variation in delay characteristics due to handover as well as the usage of contention based medium access control in some wireless link technologies.

Some of the assumptions I am explicitly making are listed below:

- There will be a need for authentication, authorization and accounting (AAA[1]) services in these types of networks.
- Authentication per packet may be necessary at link level in the access network, while privacy may only be needed at higher layers.
- A user should be able to continue ongoing sessions after a handover, at least for those real-time applications that I am considering (see section 2.3)
- There will be a need to provide fast handover support in WLANs (such support already exists within WWANs)
- In urban areas there will be full WLAN coverage, however, geographically adjacent APs will commonly be attached to *different* operators. To enable mobile users to keep WLAN connectivity while moving within these areas, these networks must be open for other operator's customers, e.g, by establishing roaming agreements.
- IP (in particular IPv6) will be used end-to-end. I will therefore ignore voice gateways to PSTNs and no transcoding of the voice data will be done along the path from the source to the receiver.
- In order to accomplish satisfactory AAA and privacy services, the existence of a Public Key Infrastructure (PKI) is assumed.
- IEEE 802.11 will be successful also for voice services (without resource reservation and without requiring the Point Coordination Function (PCF), see section 2.1.1)

1.2 Problem of concern

The main goal for this thesis study will be to achieve fast and secure handovers for mobile Internet users who are utilizing real-time services such as mobile telephony. An additional requirement is that the end-to-end delay for these services should be bounded, except perhaps just before, during, and just after a handover.

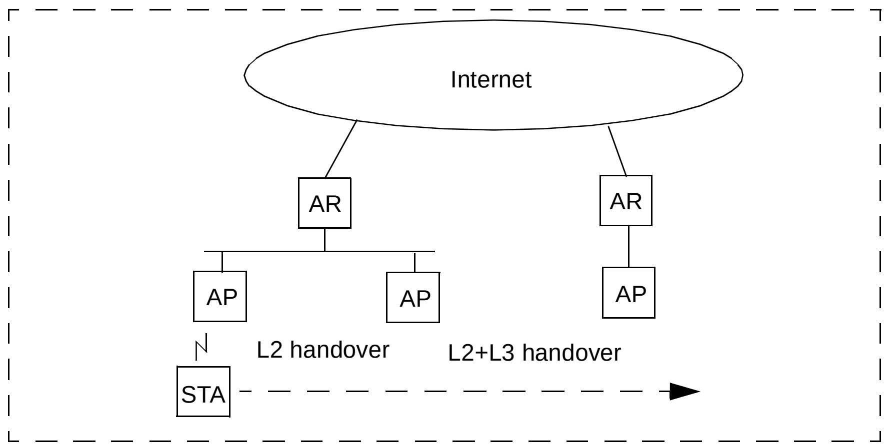
It should be noted there is ongoing work within several organizations and research groups concerning different aspects of this topic. That related work will be covered further in next chapter, however, to explain the contribution of this thesis study a brief summary is given here as well. There has been research regarding the capability of the IEEE 802.11 PHY and MAC layers to carry voice traffic (see section 2.4), there is ongoing standardization work to introduce AAA services at both layer 2 and layer 3 to enable secure access, roaming and billing (see section 2.1 and 2.2). To improve handover performance there are groups looking at context transfer both at layer 2 and 3, but possibly also at higher layers, based on the idea that handover latency will be reduced if context state (encryption keys, header compression, etc.) can be transferred instead of being re-negotiated when a mobile station (STA) moves (see section 2.1 and 2.2). There is also ongoing work to improve IP layer handover as well as work on adaptive applications (see section 2.2 and 2.3).

Although there remains work to be done within each of these sub-areas, there is little or no co-ordination between these efforts. Thus, the main contribution of this dissertation will be to take a broader approach to the problem in order to see the big picture, describe the different components involved and how they affect each other, and suggest solutions to optimize the ***overall*** *performance where possible.*

2.0 Previous work in this and related areas (literature review)

The task of achieving low handover and bounded end-to-end delay for real-time internet services could be viewed as consisting of three sub-areas: link (L2), network (L3), and higher layer issues. Figure 1 below can be used to illustrate how these sub-areas are related.

FIGURE 1. A STA performing handover first within a subnet and then across subnet borders



When a handover is made within an IP subnet, a STA is able to keep its IP address and no layer 3 mechanisms need to be involved. It should be noted that higher layer mechanisms, such as audio playout buffer schemes, may react on a layer 2 handover. The higher layer protocols may have to deal with the consequences of the layer 2 handover itself (loss of packets, etc.), but also because the traffic situation within the cell of the new AP may vary vastly from the situation within the previous cell. A handover that is made across IP subnet borders (shifting from one access router (AR) to another) will have to deal with layer 3 issues, in addition to the layer 2 and the higher layer mechanism for the layer 2 handover.

As mentioned in the previous section related work has been done concerning each of the different sub-areas of the problem; however, work that deals with the bigger picture and looks at the overall performance is rare. In this section an introduction to the relevant research within each of the different sub-areas are covered in sections 2.1-2.3. Section 2.4 covers work that spans over these sub-area borders.

2.1 Link layers

Traditionally wireless access technologies have been classified as wireless WAN (WWAN) or wireless LAN (WLAN) technologies depending on the coverage area of the radio cells. Within this study we will limit our attention to some of the WLAN and WWAN technologies, that we think are the most important for the services we anticipate. Within the WLAN market the IEEE 802.11[2] technology (in particular 802.11b) has gained a tremendous market size, and newer 802.11 technologies (802.11a and 802.11g) are expected to reach an even higher penetration in future wireless networks. A competitor to 802.11a is HyperLAN/2 standardized by ETSI. Both 802.11a and HyperLAN/2 work in the 5GHz band and operate at greater data rates than what is currently available in the 2.4 GHz band. Their major (technical) difference concerns their channel structure and access control of the medium. HyperLAN/2 is based on TDMA and resource reservation, while 802.11a is using the 802.11 medium access control and frame structure. In the following sections the focus will be on 802.11, since one of the assumptions I make is that 802.11 will become successful in carrying voice traffic.

2.1.1 IEEE 802.11 Today

To see how well IEEE 802.11 is suited to carry mobile IP telephony we are interested in both the handover procedure and its associated delay, as well as the delay properties once a STA is attached to an AP.

Regarding the delay, it is interesting to note that 802.11 defines two medium access schemes: the distributed co-ordination function (DCF) which can be referred to as a “wireless Ethernet” and the point co-ordination function (PCF) where the AP polls the STAs. The intention for defining the PCF mode was to support voice services[3], however, no AP currently on the market supports PCF. As will be described in section 2.4 there are some analytical and simulation studies regarding voice services over PCF, and the result is not very promising. DCF is a contention based scheme using CSMA/CA. The delay includes both deterministic components which depend on: current data rates, packet size, and whether WEP encryption is enabled or not (processing delay), but it also contains a random delay component inherent in contention based schemes. In [10] it is shown how the forwarding delay over a 802.11 AP can be modeled as a single server single queue system, based on measurements on two different 802.11b APs of a common vendor (Lucent). It is worth noting that the service time for the downlink was significantly higher than for the uplink (in the order of 1 millisecond and 200 microseconds respectively for a 64 byte packet; with WEP not in use).

The IEEE 802.11 standard[2] describes the messages exchanged when a STA performs a handover between two APs within the same extended service set (ESS), i.e., a re-associating; however, little work has been published regarding the handover performance in real networks[11]. When performing a handover, a STA must first authenticate to the new AP, and upon successful authentication it can issue a re-association message exchange. Note that at the link layer, a STA is only allowed to be associated with one AP at the time, thus it must disassociate with its old AP before associating with the new AP. As mentioned in [2], the authentication process can be time consuming (depending on the authentication protocol in use) and in order to improve handover performance the STA is allowed to authenticate to a new AP while still being associated with the old AP (*pre-authentication*). A somewhat simplified picture of the procedure is presented in figure 2 below.

FIGURE 2. Classic 802.11 handover

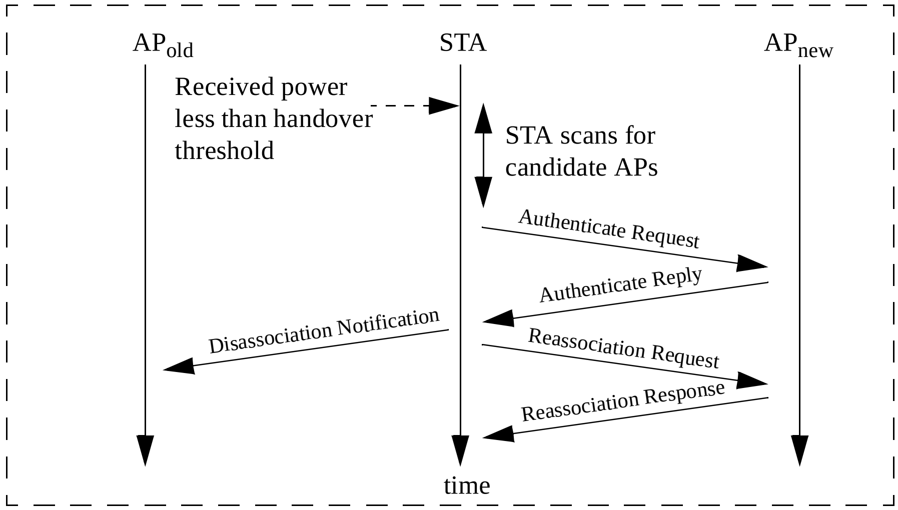


Figure 2 does not show all the details that affect the handover delay, e.g., messages for scanning for candidate APs, additional control messages (RTS, CTS, ACK) and do not show all messages between the STA and the new AP if shared key authentication is used. Nevertheless, the figure is useful when discussing handover delay. As shown in figure 2, the STA is able to pre-authenticate to a new AP while still being associated with its old AP, i.e., it can perform the authentication message exchange before sending a disassociation notification to its old AP. It would be interesting to check whether current implementations use pre-authentication and to see how that affects the handover delay, e.g., it would be of interest to see whether packets being delivered to the STA during pre-authentication or AP scanning are lost or delayed. It would also be of interest to verify the order of the messages as shown in figure 2, since a similar figure in [15 (p.8)] seems to indicate that even the association request and response message are left out of the critical path.

And as mentioned above, 802.11 defines a shared key authentication method based on the WEP algorithm, which may also affect the handover delay. Other proprietary access control mechanisms exist, such as storing lists of allowed MAC addresses in the APs or in a central (RADIUS[14]) server. WEP is also used for message encryption and integrity checking. WEP only adds 4 bytes overhead per packet; however, the per byte per packet processing increases, which may have a non-negligible affect on the end-to-end delay. What is worse is that WEP has some well-known security design flaws and it also lacks a mechanism for dynamic key distribution, which makes it unsuitable for AAA usage as well as a means to achieve privacy [6,7,8]. Therefore, initiatives has been taken both to improve and to replace WEP. The latter choice will be covered in the next section.

2.1.2 802.11, Future trends

To remedy the problems of WEP, the 802.11 Task Group I (Tgi) was given the task to define the integration/adaptation of 802.1X[9] (Network Access Port Authentication) with 802.11. 802.1X was developed to enable access control on public network ports. A port could be a physical port such as a port on an Ethernet switch, but could also be a logical port such as an 802.11 association. 802.1X enables a *supplicant* (the STA) and an *authenticator* (residing in the AP) to dynamically establish a shared secret by invoking a third party (the *authentication server*), usually with the extensible authentication protocol (EAP)[12] over RADIUS. EAP is not an encryption protocol itself, but it enables the involved parties to negotiate what security protocol to use (AES[13] is a good candidate and is likely to implemented in many products). This makes 802.1X an enabler for roaming, mapping of STA/AP associations to VPNs or VLANs, authentication and/or encryption of messages. By these features 802.1X provides good support for sharing of APs by multiple operators. There are still some security considerations for using 802.1X in and 802.11 environment[16,17], but the 802.11 Tgi is still working on these issues.

The introduction of 802.1X has some affect on the handover latency. First, it will not be possible to pre-authenticate (although a proposal of how to accomplish this was recently released[18]). The 802.1X frames are exchanged after a STA has been associated with an AP and since a STA is not allowed to be associated with multiple APs this exchange will be in the critical path. To establish the shared key between the AP and the STA this may even involve contacting a possibly distant entity (the authentication server).

In order improve handover performance the IEEE 802.11 Task group f (Tgf) is developing an inter-access point protocol (IAPP). A draft proposal[19] produced in spring 2001 was rejected at balloting, but there is ongoing work to come up with a new proposal. In the rejected draft, when a STA reassociates with an AP, the new AP will contact the old AP, requesting it to forward any stored context information about the STA. In particular, shared keys dynamically created by use of 802.1X could be transferred (when WEP is used this support is not needed[15]). The basic idea is that if context can be transferred instead of negotiated the handover delay could be reduced. But the transfer of context will require trust, and checking this will add to the delay. Quantifying the effect of context transfer versus re-negotiation would be of interest. Another issue of interest for the handover delay is that the IAAP specifies mechanisms of how to enforce the requirement that a STA can only be associated with one AP at a time.

The IAAP does not assume that all APs in an ESS are located within the same LAN. The IAAP protocol runs on top of UDP/IP, thus each AP is required to have an IP address[1]. Unfortunately IAAP is becoming more complex, the rejected draft includes the use of a registration service instead of making this a peer-to-peer protocol between the APs. One reason that a registration server is needed is to help an AP to resolve the IP address of an “old” AP when only the MAC address of that AP is known. Even if a registration server is useful, organizations may be unwilling to introduce yet another server in their network, which in the end may limit the success for the IAPP protocol. IAPP as formulated in the draft was designed so that APs use IPv4 address of a common address space, i.e., how to handle IPv6 or deal with NAT gateways were not specified. These and other issues regarding the use of IAPP and 802.1X together with 802.11 are discussed in [20].

1. This IP address could be in a private IP address space, just as most routers/bridges within an ISP are.

2.1.3 WWAN

L2 handover within wireless LAN is handled within the cellular system and is not dealt with within this dissertation. What is of interest for me is to understand matters concerning the L3 handovers such as the establishment and tear down of WWAN connections. In addition to that the authentication and sending time lines for GSM (data channel) and GPRS should be studied to evaluate their ability to handle IP telephony. Third generation WWAN systems, such as UMTS and CDMA2000, will not be covered within this dissertation.

2.2 IP layer mobility

When companies and universities provides wireless LAN access to their employees, the APs are commonly hooked up on the same LAN. In such networks, a STA can commonly roam within a geographical area (such as a campus) while only performing layer-2 handovers. However, there are situations when a STA after a handover ends up on another subnet. One example would be if a STA moves within urban areas where different house owners, companies or ordinary operators have deployed WLAN access, but no-one provides full coverage. Another case would be if STA is moving in and out of areas with WLAN coverage, thus switching between WLAN and WWAN access. There may also be cases where roaming between APs belonging to the same operators may result in a layer-3 handover, since operators may of various reason like to subnet their network. When a handover is performed between IP subnets, routing as well as AAA[22] and context transfer issues may affect the end-to-end and handover delay.

The routing issues relates to the need for a STA to acquire a new care-of IP address when attaching to a new subnet. The downstream and upstream packet flows must be redirected, and meanwhile packets may be lost. Furthermore, in many IP based mobility schemes the data packets traverse a non-optimal path via a “home network”, which adds to the end-to-end delay. In my licentiate thesis I have looked at how the mechanisms of acquiring a care-of IP address as well as the how the characteristics of different IP based mobility schemes affect the handover and end-to-end delay[24]. There is also ongoing work within IETF Mobile IP (MIP) working group[25] for enhancing layer-3 handovers[26,27[1],28]. It should be noted that some of those suggestions rely on specific layer-2 support in order to perform well, however, it is not within the scope of those suggestions to specify how (and if) these layer-2 requirements could be fulfilled. There have also been studies of how to use MIP to handle handover between GPRS and WLAN networks[29, 30].

Depending on how the access control is implemented, a STA may have to take additional AAA related actions when moving to a new IP subnet. There are two different issues of concern: the first is to make sure that the routing redirection signalling, i.e., the binding updates, is performed in both a scalable and a secure way, and the second is to deal with the access control mechanism (above layer 2) in operation within at the new subnet. Regarding the former case, I am co-advising a MSc. degree project evaluating different proposals suggested within the IETF. Regarding the second issue, it should be noted that the handover delay can be affected by higher layer AAA solutions; today it is popular to use web-based login applications to enforce access control[23,31]. The two

1. This draft was an individual submission and not made within the Mobile IP WG track.

most significant advantages of using web-based login are that the procedure is simple enough for an ordinary user to get started and that no additional software needs to be installed on the client (a web client with support for HTTPS is usually sufficient). However, web-based login requires the user to manually enter login name and password, and doing that each time a user moves to a new subnet would not be acceptable for mobile telephony. Anyhow, web-based login is likely to exist for a long period of time in most WISP networks, so solutions to handle access control for mobile telephony will have to co-exist with web-based login methods. They could either be independent alternatives, but there might also be a possibility to make partly integrate them. e.g., you could have a common server interface, but use either an ordinary web interface or some script at the client to send the data automatically and thereby avoiding unwanted manual interaction. Security would be an issue, and knowing when to trigger such a script to run and how to adapt it to the proprietary web-login systems used by different operators would be a tricky (but probably feasible) task.

The IETF Seamoby working group[32] is investigating different ways to improve handover performance by transfer of layer 3 (or higher) context information. The basic idea is the same as for the IAAP protocol; if context can be transferred instead of negotiated the handover delay could be reduced. Examples of layer 3 context could be information related AAA, QoS and header compression[34]. The Seamoby group is also looking at the mechanisms of how to transfer such context information, e.g., they are looking at suitable ways to discover candidate access routers (CARs) in advance of a handover[33] and the design of the protocol to be used between geographically adjacent access routers (GAARs)[35]. What context parameters to transfer and how to represent them remains to be standardized, but it would clearly be of interest to evaluate and quantify the possible benefit of transferring context versus re-negotiation.

2.3 Higher layers

At the link and network layers we are interested in the mechanisms that affect the handover or end-to-end delay, while at higher layers we are mainly concerned with the applications that have to deal with the resulting connectivity. In my dissertation I will focus on real-time services and in particular telephony services (VoIP).

2.3.1 Telephony services in packet in IP networks

To provide telephony services in an IP network two main alternatives exist, H.323[37] and SIP[36]. SIP is a signalling protocol used to initiate sessions which then exchange various forms of content, although here we are mostly interested in using SIP to manage voice calls. A good overview of SIP and the other involved protocols (SDP, RTP etc.) are given in [38].

To cope with the variation in end-to-end delay in a packet network, the end-system uses a playout buffer where packets are kept until a certain playout time. Packets that are received in time can be playout in a timely manner while packets that arrive after the playout point are considered lost. The playout point is set with respect to estimated end-to-end delay and variation. There is a trade-off when setting the playout delay between packet loss due to late arrival and decreased user quality due to additional delay. Therefore, it may be advantageous to *adapt* the playout point to changes in delay and delay variation conditions [39,40,41]. Adaptive playout algorithms commonly use a fix playout point for each talk spurt, and the adapt the playout point by extending or reducing the silence periods between the talk spurts. [39] gives a comparison of different algorithms and an algorithm suitable to handle delay spikes is presented. While performing a handover, the delay and delay variation characteristics may change as a “step function”, and it remains to be tested if that proposed algorithm works fine for this case as well, and to find out how fast the algorithm converges.

The user perceived quality is also dependent on the encoding scheme in use (e.g., the packetization delay at the sender adds to the end-to-end delay) and even the CODEC in use (e.g., some codecs may cope better with packet loss than others). Thus, it may be advantageous to shift to another audio encoding scheme depending on the current network conditions. [42] presents a formula[1] for the total user perceived quality, which could be used to state what combination of buffer delay, encoding scheme, encoder, etc. to use under given conditions.

The processing at the end-nodes and its effect on the end-to-end delay should not be neglected. [41] shows how important efficient end-system application implementation and operating system support are for the end-result.

2.3.2 Other applications affected by handover

There are other real-time services which depend on high performance handovers, e.g., (interactive) video applications and virtual reality games for mobile users. Just as for voice, video data is buffered at the receiver to cope with the delay jitter caused by the network. Video applications differ from voice applications in the sense that the amount of data is larger. To achieve synchronous voice and audio during a video conference the same playout timer should be used for both the video an the audio stream. In general it is more important for a user that the audio samples are played out timely than for the video frames to be displayed timely. If video packets are lost due to handover, or dropped due to late arrival, the previous frame could be displayed again.

In virtual reality games the involved players exchange event data. Loss of data could be problem, since it would lead inconsistent state at the sites for the different players, however, virtual reality applications are generally designed to cope with that.

2.4 Cross-topic research

In the previous section the different system components were presented together with ongoing advances within each of these topics. In this section relevant research that looks at a bigger picture is presented.

2.4.1 Telephony over WLAN

As mentioned in section 2.1.1, 802.11 also has polling based MAC scheme (PCF) intended to serve real-time applications such as telephony[3]. However, PCF has not been well received by the AP vendors and PCF is not used today except in some special products. One of the problems with the PCF mode is that the 802.11 does not define how to use it, so development of compatible products is difficult. Another problem is that PCF despite its intention to serve telephony services may not be suitable for this purpose. There have been some simulation studies of the use of PCF for telephony [4,3],

1. Their formula is a based on the ITU E-model[43] with some simplifications in order to make the calculations less processing intense.

but their results, e.g. regarding delay bounds, does in my opinion not talk in favor of PCF for telephony. The IEEE 802.11 Task Group E (Tge) are working on developments in this area, and the outcome may be a new MAC scheme known as the enhanced distributed coordination function (EDCF). [5] presents a simulation study comparing a draft EDCF proposal with PCF and two other interesting MAC schemes, and shows that EDCF is an improvement upon PCF.

However, it is questionable if service differentiation within 802.11 is the right way to go. It is worth noting that Symbol Technology already provides a voice handset for 802.11[44].

2.4.2 Handover and playout buffer adaptation

As mentioned in section 2.3.1 there has there has been studies of how different buffer adaptation algorithms are able to handle delay spikes [39], but it remains to be studied whether these algorithms are also able to cope with “step” changes in delay characteristics and how fast they then can converge.

[46] deals resource reservation in 802.11 networks, but also looks adaptive playout buffer schemes from the handover perspective. The study has a quite limited scope and is not very thorough in this respect, e.g., everything except the delay involved for establishing resource reservations ignored.

2.4.3 General Mobile VoIP research

There have been some general studies on providing VoIP for mobile users, e.g. [38,45,47]. These studies gives a good overview of the protocols involved and the issues to deal with, but does provide any thorough information about handover and end-to-end delay results.

2.4.4 Combining layer 2 and layer 3 context transfer

Instead of performing layer 2 and layer 3 contexts separately, it might be possible to reduce the handover delay by integrating these solutions. It might also be desirable to perform layer 2 context transfers while performing a handover between different types of link technologies. Current work does not talk in favor of this (see below), but it would still be worth looking into.

[48] discuss how IAPP could be used to transfer AAA related context information in a 802.1X environment when performing handover using the same or different link technologies, and within and across administrative domains. Central to the model is the notion of a “correct” context transfer -- a transfer resulting in the same context on the new AP as would have resulted had a AAA conversation been completed. The conclusion is that context transfer is most likely to be successful within a homogeneous device deployment within a single administrative domain.

[34] discourages the adoption of a layer-2 protocol as an optimization measure for layer-3 context transfers, since “seamless mobility of a mobile host having Layer 2 network interfaces that support multiple radio protocols would be difficult to achieve.”

2.4.5 Combining layer 2 and layer 3 authentication

[50] presents a scheme where MIP authentication (layer 3) is integrated with 802.1X (layer 2). This is very interesting and should be studied further.

3.0 The candidate's ideas and preliminary results

In the previous chapter the different components needed to provide real-time services to mobile Internet users. My goal for the dissertation is to get a clear understanding of the full path of events that adds to the handover and end-to-end delays for mobile Internet users utilizing real-time services such as VoIP. In order to do so I would need to acquire quantitative values of each component as well as on the total delay for some candidate solutions. For some of these components I intend to acquire these values by practical measurements, while others may be based analytical evaluation and possibly even simulations.

Where possible, I will suggest enhancements in order to increase the performance or e.g., to adapt the systems to IPv6. As there will be several design alternatives related to the individual components as the system as a whole, an identification of critical properties (key resources etc.) within the suggested solutions should be identified. The reason for this is to compare the different alternatives not only from an end-to-end and handover point of view, but also from their probability of successful deployment.

3.1 Handover performance

Below I have listed some of the main questions to look at to understand and quantify the handover delay for mobile VoIPv6:

- What delay is introduced by the different components?
  - Layer-2 handover using the existing and coming alternatives described in section 2.1 (WEP, 802.1X and IAPP).
  - Layer-3 handover, with main focus on the causes of L3 context switching, enhanced mobile routing support and AAA issues (less focus on fast handover between WLAN and WWAN networks)
  - Convergence delay caused by higher layer applications, in particular playout buffer and encoding adaptations.
- How do the delay introduced by these components affect the total handover delay?
  - Are the delays additive or can some of them be masked by doing things in parallel or in advance? An example would be to study the trade-off between handover delay and operator revenue if allowing/refusing traffic to pass before the user has been properly authenticated.
  - Would it be possible to utilize layer 2 information to trigger events at layer 3 or higher, without losing the desired generality of the layered model?
- What is the system's robustness to packet loss?
  - How could the handover be affected in case of the loss of single packet as compared to the best case when no packets are lost?

3.2 End-to-end delay

The main questions to be answered to understand and quantify the end-to-end delay are the following:

- How does the voice encoding, mobility support scheme, and wireless link technology (802.11 and GPRS/GSM) add to the overall end-to-end delay? The delay caused by the backbone network should also be taken into account, but no detailed study will be performed on that topic.
- What will is the average versus the worst case delay, and how often does non-average behavior happen?
- Is it possible to achieve acceptable quality over a 802.11 WLAN without using any type of resource reservation scheme?

3.3 Key resources, complexity etc.

The solution to support real-time services should not only have high performance, but must also be deployable. In order to compare the different candidate solutions from this aspect, a rudimentary evaluation framework should be established and key resources should be identified. Examples of aspects that could be listed in this framework could be:

- Processing demand at critical points in the system, e.g., encryption and decryption within APs.
- Data storage and transfer, e.g., signalling and storage of accounting information
- Complexity

4.0 Action / Time plan / Milestones

For this study a set of different activities have been identified. They will be described below. Following this comes a preliminary schedule for these activities.

- Initial literature survey

  An initial literature survey is part of this dissertation proposal. This literature survey will be regularly updated, as I will come across more references when digging deeper into the different topics.

- Dissertation proposal

  The proposal (the document you are currently reading) should be presented and defended at an official seminar.

  Deliverable: A dissertation proposal document and a seminar

  Date: Seminar June 5

- Study on handover delay within 802.11 distribution system (DS)

  The study of the delay related to handovers within an 802.11 distribution system will be divided into three different parts: 802.11 as used today (WEP), 802.11 combined with 802.1X, and finally 802.11 with 802.1X and IAPP. Each part will be studied both theoretically and practically. To do practical tests both a test network and suitable measurement methods have to be established. To establish a test network with support 802.1X and 802.11f I intend to initiate and supervise two M.Sc. projects with the aim to implement these protocols. Regarding measurement methods there is a need to develop/find tools or functionality to detect when a handover has occurred, and it would also be desirable to have tools to trigger a handover to occur (handover coercion[51]).

  Deliverables:

  - A theoretical study on 802.11 handover, finished July 15
  - Handover trigger and detection functionality, finished July 15
  - A practical study on 802.11 handover, finished July 15
  - A theoretical study on 802.1, 802.1X and 802.1f handover, finished July 30
  - A practical study on 802.1, 802.1X and 802.1f handover, finished Nov. 5

  While performing these studies the associated end-to-end delay should also be studied.

- Layer 3 mobility studies:

  One of the main objectives for the layer 3 mobility studies is to study the traffic pattern (packet loss, delay and arrival order, ...) during a handover, since this what the VoIP applications need to deal handle. Some key components include

  - Routing issues, i.e., packet delay, packet loss and arrival order for Mobile IPv6.
  - AAA issues, additional handover delay introduced by AAA services
  - Context transfer (seamoby)

Deliverable: A study on handover and end-to-end delays, based on both analytical and practical measurements. Finished Sep. 25.

- Voice over Mobile IPv6

  What is interesting to study is *if* and *how* mobility affects voice over IP services. Of particular interest is the possible effect of handover on the playout buffer and choice of encoding schemes. Some initial tests will be finished Oct. 30.

- VoIP, Layer-3 and Layer-2 (802.11)

  The most interesting part will be to put all of it together and measure the overall performance. Deliverables:

  - An analytical and practical evaluation of handover and end-to-end delay for voice applications when performing L2, and L2 plus L3 handovers in an 802.11 environment supporting 802.1X and IAPP. Finished Jan 20, 2003.

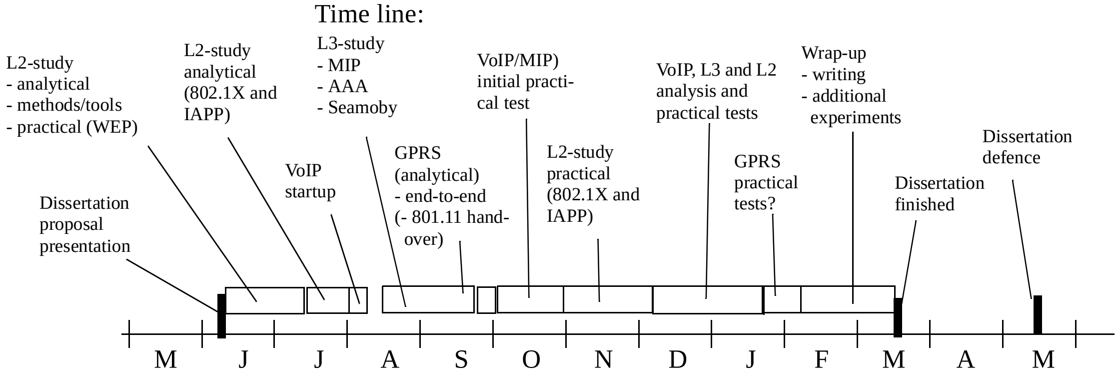
- GPRS

  To gain understanding of end-to-end delay in GPRS networks, both a theoretical and some practical measurements will be performed. In case some co-operation with Ericsson can be established, it may also be possible to perform practical studies on 802.11 to GPRS handovers.

  Deliverables:

  - An analytical study of end-to-end delay in GPRS networks. Finished Oct. 1.
  - A practical measurements on end-to-end delay in GPRS networks and possibly also on GPRS/WLAN handover. Finished Feb. 5, 2003.

Time line:



5.0 Rough outline of the thesis

1 Introduction

1.1 Handover delay overview

1.2 End-to-end delay for Mobile VoIP

2 Wireless systems

2.1 802.11

2.2 GPRS

3 IP mobility

4 Voice over IP

5 Putting it all together

5.1 Total system design alternative examples

5.2 Evaluation of alternatives

6 Conclusions and future work

6.0 References

[1] Bernard Aboba et.al. RFC 2989, “Criteria for Evaluating AAA Protocols for Network Access“, Nov. 2000

[2] IEEE Std. 802.11, 1999 Edition (ISO/IEC 8802-11:1999) IEEE Standard for Information Technology -Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Network -Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications

[3] Al Petrick, Jim Zyren, Juan Figueroa, “Delivering Voice over IEEE 802.11 WLAN Networks“, Harris Semiconductor http://www.intersil.com/design/prism/papers/Voice_%20over_IEEE_802_11.htm

[4] Malathi Veeraraghavan, Nabeel Cocker, Tim Moors, “Support for voice services in IEEE 802.11 wireless LANs”, Proceedings of Infocom 2001, pp. 488-97, Apr. 2001

[5] Anders Lindgren, Andreas Almquist and Olov Schelén, “Evaluation of Quality of Service Schemes for IEEE 802.11 Wireless LANs” In Proceedings of the 26th Annual IEEE Conference on Local Computer Networks (LCN 2001), November 15-16, 2001 Tampa, Florida, USA. http://www.sm.luth.se/~dugdale/publications/lcn2001lindgren.pdf

[6] Jesse R. Walker, “Unsafe at any key size; An analysis of the WEP encapsulation”, Oct. 27, 2000, http://www.drizzle.com/~aboba/IEEE/0-362.zip, Accessed March 2002

[7] Nikita Borisov, Ian Goldberg, David Wagner, “Intercepting Mobile Communications: The Insecurity of 802.11”, and/or http://www.isaac.cs.berkeley.edu/isaac/wep-faq.html, Published in the proceedings of the Seventh Annual International Conference on Mobile Computing And Networking, July 16–21, 2001

[8] W. A. Arbaugh, N. Shankar, and Y. J. Wan. Your 802.11 wireless network has no clothes. http://www.cs.umd.edu/waa/wireless.pdf, Mar. 2001.

[9] IEEE Std. 802.1X-2001, IEEE Standard for Local and metropolitan area networks: Port-Based Network Access Control

[10] Iyad Al Khatib, Rassul Ayani and Gerald Q. Maguire Jr., “Wireless Access Points as Queuing Systems: Delay Probability and Service Time”, Submitted for publication

[11] Mobile IP Mailing list archive, E-mail by James Kempf in the discussion “IPv4 Fast Handover Support for 802.11” http://playground.sun.com/mobile-ip/WG-archive/msg05737.html, Accessed May 2002.

[12] Larry Blunk, John Vollbrecht and Bernard Aboba, “Extensible Authentication Protocol (EAP)”, IETF draft <draft-ietf-pppext-rfc2284bis-04.txt>, April 2002, work in progress.

[13] Advanced Encryption Standard (AES) home page, http://csrc.nist.gov/encryption/aes/, Accessed May 2002.

[14] Carl Rigney, Allan Rubens, William Simpson and Steve Willens, RFC 2865 “Remote Authentication Dial In User Service (RADIUS)”, June 2002

[15] Tim Moore, Bernard Aboba, “Authenticated Fast Handoff”, http://www.drizzle.com/~aboba/IEEE/11-01-TBD-I-Authenticated-FastHandoff.ppt, Accessed May 2002.

[16] Bernard Aboba, “The Unofficial 802.11 Security Web Page”, http://www.drizzle.com/~aboba/IEEE, Accessed May 2002.

[17] Arunesh Mishra, William A. Arbaugh, “An Initial Security Analysis of the IEEE 802.1X Standard”, Technical Report CS-TR-4328 UMIACS-TR-2002-10, University of Maryland, 6 Feb 2002. http://www.cs.umd.edu/~waa/1x.pdf

[18] Bernard Aboba, “802.1X pre-authentication”, submission to IEEE 802.11 Tgi, May 14, 2002, http://www.drizzle.com/~aboba/IEEE/11-02-TBDr0-I-Pre-Authentication.doc

[19] IEEE P802.11 Wireless LANs, IEEE 802.11f pre-Draft P802.11f/D0.2 “Recommended Practice for Multi-Vendor Access Point Interoperability via an Inter-Access Point Protocol Across Distribution Systems Supporting IEEE 802.11 Operation”, March 14, 2001

[20] John Vollbrecht, David Rago, Robert Moskowitz,” Wireless LAN Access Control and Authentication”, White Paper, Interlink Networks, http://www.interlinknetworks.com/references/papers/WLAN_Access_Control.pdf, Accessed May 2002.

[21] ETSI 3rd Generation Partnership Project (3GPP), “Universal Mobile Telecommunications System (UMTS); Handovers for real-time services from PS domain”, 3GPP TR 25.936 version 4.0.1 Release 4, Dec. 2001

[22] Steven Glass et. al., RFC 2977 “Mobile IP Authentication, Authorization, and Accounting Requirements”, Oct. 2000

[23] Martin Hedenfalk, “Access Control in an Operator Neutral Public Access Network“, Master Thesis, Dept. of Microelectronics and Information Technology, Royal Institute of Technology, Sweden, http://www.e.kth.se/~e97_mhe/thesis/thesis.pdf

[24] Jon-Olov Vatn, “Improving Mobile IP handover performance”, Licentiate thesis, June 2002, TRITA-IT AVH 00:06, ISSN 1403-5286, ISRN KTH/IT/AVH--00/06--SE, Department of Teleinformatics, Royal Institute of Technology, Stockholm, Sweden

[25] IETF Working group on “IP Routing for Wireless/Mobile Hosts (mobileip) “, http://www.ietf.org/html.charters/mobileip-charter.html.

[26] Alper Yegin et.al “Fast Handovers for Mobile IPv6”, IETF draft <draft-ietf-mobileip-fast-mipv6-04.txt>, Work in progress, March 2002.

[27] Karim El Malki and Hesham Soliman, “Simultaneous Bindings for Mobile IPv6 Fast Handoffs”, IETF draft <draft-elmalki-mobileip-bicasting-v6-01.txt>, Work in progress, Nov. 2001.

[28] Hesham Soliman et. al., “Hierarchical MIPv6 mobility management (HMIPv6)“, IETF draft <draft-ietf-mobileip-hmipv6-05.txt>, Work in progress, July 2001

[29] Jonas Mohlin, “Mobile IP in GPRS and IEEE 802.11 wireless LAN“, M.Sc. thesis project, Department of Microelectronics and Information Technology, Royal Institute of Technology, Sweden, May 2002, http://www.e.kth.se/~e97_jmo/xjob/files/report.pdf)

[30] Kaveh Pahlavan, Prashant Krishnamurty, Ahmad Hatami, Mika Ylianttila, Juha-Pekka Makela, Roman Pichna, Jari Vallström, “Handoff in Hybrid Mobile Data Networks”, IEEE Personal Communications, pp. 34-47, April 2000

[31] Juan Caballero and Daniel Malmkvist, “Experimental Study of a Network Access Server for a public WLAN access network”, M.Sc. thesis project, Department of

Microelectronics and Information Technology, Royal Institute of Technology, Sweden, Jan. 2002, http://www.e.kth.se/~e97_dma/FinalReport.pdf

[32] IETF Working group on “Context Transfer, Handoff Candidate Discovery, and Dormant Mode Host Alerting (seamoby)”, http://www.ietf.org/html.charters/seamoby-charter.html. Accessed May 2002.

[33] Dirk Trossen, Govind Krishnamurthi, Hemant Chaskar and James Kempf, “Issues in candidate access router discovery for seamless IP-level handoffs”, IETF draft <draft-ietf-seamoby-cardiscovery-issues-02.txt>, work in progress, 15 January 2002

[34] James Kempf, “Problem Description: Reasons For Performing Context Transfers Between Nodes in an IP Access Network”, IETF draft <draft-ietf-seamoby-context-transfer-problem-stat-04.txt>, Work in progress

[35] Hamid Syed, Gary Kenward, Pat Calhoun, Madjid Nakhjiri, Rajeev Koodli, Kulwinder Atwal, Mark Smith and Govind Krishnamurthi, “General Requirements for Context Transfer” IETF draft <draft-ietf-seamoby-ct-reqs-03.txt>, Work in progress, January, 2002

[36] Mark Handley, Henning Schulzrinne, Eve Schooler and Jonathan Rosenberg, IETF RFC 2543 “SIP: Session Initiation Protocol”, Mar. 1999

[37] ITU-T, “Recommendation H.323, Packet-based multimedia communications systems”, Oct. 1997

[38] Henning Schulzrinne and Jonathan Rosenberg. The IETF Internet Telephony Architecture and Protocols. IEEE Network, pages 18-23, May/June 1999

[39] Ramachandran Ramjee, Jim Kurose, Don Towsley, and Henning Schulzrinne. Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks. In IEEE Infocom, pages 680-688, 1994.

[40] Sue B. Moon, Jim Kurose, and Don Towsley, “Packet Audio Playout Delay Adjustment: Performance Bounds and Algorithms”, ACM/Springer Multimedia Systems, Vol. 6, pp. 17-28, January, 1998

[41] Ian Marsh, Olof Hagsand and Kjell Hanson, “End system adaption to packet audio jitter”, Submitted to ACM Multimedia 2002

[42] R.G Cole and J.H. Rosenbluth. Voice over IP Performance Monitoring. Computer Communication Review, ACM Sigcomm, 31(2):9-24, April 2001.

[43] ITU-T recommendation G.107: The E-Model, a computational model for use in transmission planning

[44] Symbol Technologies’ Netvision phone, http://www.symbol.com/products/wireless/nv_phone.html

[45] Prathima Agrawal, Jyh-Cheng Chen, Cormac J. Sreenan, “Challenges for Mobile Voice-over-IP” MWCN 2000, LNCS 1818, pp.58-69, 2000

[46] Andreas Terzis, Mani Srivastava, Lixia Zhang, Simple QoS Signaling Protocol for Mobile Hosts in the Integrated Services Internet, Proc. of IEEE INFOCOM’99, March 1999

[47] Theo Kanter, Christain Olrog, Gerald Q. Maguire Jr., VoIP over Wireless for Mobile Multimedia Applications, Proceedings of the 1999 Personal Computing and Communications Workshop,pp. 12-19, Nov 1999

[48] Bernard Aboba, Tim Moore, A Model for Context Transfer in IEEE 802, IETF draft <draft-aboba-802-context-01.txt>, Work in progress, 28 August 2001

[49] Stephen Kent and Randall Atkinson, RFC 2401 "Security Architecture for the Internet Protocol", Nov. 1998

[50] Luca Salgarelli et. al., "EAP SKE authentication and key exchange protocol" IETF draft <draft-salgarelli-pppext-eap-ske-01.txt>, work in progress, April 2002

[51] A. Helal, C. Lee, Y. Zhang, and G. Richard III. An Architecture for Wireless LAN/ WAN Integration. In IEEE Wireless Communications and Networking Conference (WCNC 2000), pages 1035-1041, sep 2000.

7.0 Acronyms and abbreviations

AAA	Authentication, Authorization and Accounting
AES	Advanced Encryption Standard
AP	Access Point
AR	Access Router
BSS	Basic Service Set
CAR	Candidate Access Router
CSMA/CA	Carrier Sense Multiple Access / Collision Avoidance
CODEC	Coder/Decoder
DCF	Distributed Coordination Function
EAP	Extensible Authentication Protocol
EDCF	Enhanced Distributed Coordination Function
ESS	Extended Service Set
ETSI	European Telecommunications Standards Institute
PHY	Physical Layer
GAAR	Geographically Adjacent Access Router
GSM	Global System for Mobile communications
GPRS	General Packet Radio Service
HTTPS	Hypertext Transfer Protocol Secure
IAPP	Inter-Access Point Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
ISP	Internet Service Provider
MAC	Medium Access Control
MIP	Mobile IP
PCF	Point Coordination Function
PKI	Public Key Infrastructure
PSTN	Public Switched Telephony Network
RADIUS	Remote Access Dialin User Service
RTP	Real-time Transport Protocol
SDP	Session Description Protocol
SIP	Session Initiation Protocol
STA	Mobile Station
TDMA	Time Division Multiple Access
VoIP	Voice over IP
WEP	Wired Equivalent Privacy
WISP	Wireless Internet Service Provider
WLAN	Wireless Local Area Network
WWAN	Wireless Wide Area Network