Network-controlled mobility within radio access networks based on WLAN technologies
Une gestion de la mobilité contrôlée par le réseau dans des réseaux d'accès IP basés sur la technologie RLAN

Authors:
Karine Guillouard, Philippe Bertin, Youssef Khouaja, Jean-Christophe Rault, France Telecom R&D
Jean-Marie Bonnin, ENST Bretagne

Abstract

This article presents a network-controlled approach of user terminal mobility within an IP based Wireless LAN Access Network.
In a first part, this article makes a review of the mobility support, on the subject of emerging WLAN technologies as HiperLAN/2 and IEEE 802.11, on the one hand, and, regarding IP networks as currently studied within IETF, on the other hand. Both types of IP mobility protocols are presented, either global mobility protocols such as Mobile IP, or local mobility management protocols (micro mobility).
In the next part, the overall principles of our mobility management approach are explained; this approach is based on the implementation of a new network entity dedicated to the control of user terminal mobility.
The last part details a practical implementation of this approach. The implementation is carried out on the basis of Hierarchical Mobile IPv6 (HMIPv6). The experimental results confirm the importance to carefully plan and control the user terminal mobility within large IP based Access Networks, as this brings benefit to the user as well as to the operator.

Résumé

Cet article présente une approche de gestion, contrôlée par le réseau, de la mobilité des terminaux dans un réseau d'accès IP construit à partir de réseaux radio locaux haut-débit (RLAN). Avant d'aborder les principes et les intérêts de cette approche, cet article dresse un rapide état de l'art du support de la mobilité, d'une part, dans les RLAN émergents de type IEEE 802.11 et HiperLAN 2, et d'autre part, dans les réseaux IP en général, tel que cela est suggéré par les actuelles propositions à l'IETF de protocoles de mobilité IP. Les protocoles IP gérant la mobilité globale, tel que Mobile IP, aussi bien que la mobilité locale (micro-mobilité), sont alors étudiés. Ensuite, les grandes lignes de notre approche de gestion de mobilité, basée sur la mise en œuvre dans le réseau d'une entité capable de contrôler la mobilité des terminaux, sont expliquées. Le papier termine en détaillant une implémentation de cette approche, réalisée dans les laboratoires de France Télécom et basée sur le protocole HMIPv6 (Hierarchical Mobile IPv6). Les résultats expérimentaux obtenus confirment l'intérêt, aussi bien pour l'opérateur que pour l'utilisateur, de planifier et de contrôler par le réseau le déplacement des terminaux dans un réseau d'accès IP de grande envergure.

Glossary

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AP</td>
<td>Access Point</td>
</tr>
<tr>
<td>APC</td>
<td>Access Point Controller</td>
</tr>
<tr>
<td>APT</td>
<td>Access Point Transmitter</td>
</tr>
<tr>
<td>AR</td>
<td>Access Router</td>
</tr>
</tbody>
</table>
1 Introduction

Most European network operators have already introduced or intend to introduce in a very near future their GPRS network, which is commonly known as 2.5G. This naming implies that GPRS is generally perceived as an intermediate step towards the implementation of 3rd generation networks; actually, GPRS is already capable of providing additional features to the user, as higher bit rates, packet based charging and always on connections. The upcoming 3G systems (UMTS) will improve one step further, by providing even higher data rates, as well as improved option regarding quality of service and bandwidth reservation.

At the moment, a remarkable interest for local radio access network (Wireless LAN) can be observed. WLAN are used in several situations, which can cover the need of individuals such as ADSL extension, as well as companies using them to allow a flexible access to their intranet. WLAN are also appropriate to cover temporary events such as forums and conferences. Lastly, operators plan to use them as systems able to provide Internet access or corporate access in public or private locations that are called “hot-spots”, which include airport, railway stations and hotels for example.

In the actual situation, WLAN and UMTS can be considered as complementary systems; WLAN are more dedicated to computer services, dealing with data in the same way as an Ethernet network; UMTS is assigned the task of dealing with real time voice applications and multimedia data. In fact, a number of network operators have already started to make use of these complementary features and thus to deploy WLAN, but both systems are generally unconnected and there are few possibilities for the users to freely roam from one network to the other.

Under these conditions, and even if it is always a difficult task to predict the long term situation, especially in the mobile telecommunication field, it can be assessed that the next generation of mobile network, currently designated by “beyond 3G”, will be built to complement the UMTS capabilities, both regarding the available bandwidth and the ability to handle non symmetric communications. These improvements aim at supporting broadband multimedia applications (including high quality video flows that can request data rates higher than 2Mbit/s), as well as services like web surfing demanding far more downlink than uplink.
capacity. The other main feature of “beyond 3G” systems will probably address the variety of access networks, which will allow the user to access to his services in a transparent way; actually the various access networks will be cooperating and thanks to his multi-homed terminal the user will not need to care about which network provides which services.

Even if the whole set of actors involved in this field have not yet come to an agreement, several ideas are mentioned very regularly when addressing the “beyond 3G” issue, as listed below:
- Use of an all IP network that is QoS enabled
- Addition of “micro-cellular” technologies such as Wireless LAN to complement UMTS
- Implementation of multi-interface user terminals, capable of handling vertical handovers between access networks based on different technologies.

Several European entities are actively involved in “beyond 3G” architecture with respect to the standardization process as well as to more prospective and research issues addressed for example by IST projects funded by the European Commission.
A standardization group at ETSI BRAN actually works on the interworking between WLAN (initially HiperLAN type 2) and 3G networks. Reference architecture has been defined, and technical detailed aspects are still under discussion, namely regarding accounting and billing. This ETSI BRAN group has joined with the Japanese MMAC group currently in charge of 5GHz WLAN, and is under negotiation with IEEE, with the aim of agreeing on a true global standard.

On the other hand, within the framework of IST projects as BRAIN [29] and MIND [28], a number of concepts have been studied aiming at progressing towards “beyond 3G” systems. In that field, a significant amount of work has been dedicated to figuring out how to adapt WLAN to make them more efficient regarding IP packet transport, which may support real time applications. For that purpose, IP access networks based on WLAN technologies will be requested to offer new features such as quality of service and mobility management, so that broadband real time multimedia application can be properly supported when accessed by a mobile terminal that will likely perform handovers during a session.

This short introduction has presented our view of the “beyond 3G” situation, and the future significant role of WLAN. In this context, the present paper will focus on WLAN and more particularly on the way to handle the mobility within WLAN.

The paper will start dealing with the different issues related to mobility, by reviewing the state of the art of mobility handling within WLAN in a first stage, as well as mobility at the IP level in a second stage.
The next section of the paper will address the concepts studied and developed within France Telecom research laboratories; actually, these concepts concentrate on mobility management schemes that are network-controlled within wireless IP access networks.
The last part of the paper will focus on the performances that are reached when using a network controlled mobility management scheme. Handover delays will be presented, leaning on theoretical and simulation work. In addition a test platform implementing an IP access network based on a WLAN is described. The implementation of the network-controlled IP mobility has allowed us to perform real time measurements, which are able to confirm the results expected from the theoretical and simulation work.
2 Mobility in WLAN

As introduced above, the wireless LAN technology can be thought of as a complementary access technology for the design of all IP cellular access networks. The purpose is to extend the coverage of 3G cellular networks with hot-spots providing locally a higher bandwidth. However the support of mobility in wireless LAN is quite different compared to cellular networks. Globally, wireless LAN offer:

- lower-scale mobility due to a limited area coverage and to the support of roaming only between access points of a same technology and belonging to the same LAN (see Figure 1);
- slower motion (e.g. pedestrian speed), even if current studies, within IST project for example, aim at allowing faster motion (e.g. car or train speed);
- smaller range about 20-40m for indoor areas and 100-200m in open space to be compared to kilometres for cellular systems; the short range is due to a limited transmission power for practical (power consumption saving) and regulatory reasons (limited EIRP in WLAN bands, i.e. at 2.45GHz and 5GHz); the range of a radio cell covered by an access point is directly tied to the overall required bandwidth within the radio cell;
- and more generally, no WAN (Wide Area Network) features are supported (scalability, address management, localisation, customer information management): such features belong to higher layers that are not addressed by wireless LAN standards.

Concerning the mobility management, the way to handle handoff between access points of a WLAN globally depends on the configuration of these access points. If the access points are localised in the same IP subnet, the handoff is directly managed by link level protocols between the involved MT (Mobile Terminal) and APs (Access Points). No third entity is required to re-associate the MT and to maintain the active connections. On the contrary, if the APs belong to different IP subnets, higher level protocols are also to be used to manage the change of location of the MT. For example, DHCP [8] may be used in order to allocate a new IP address corresponding to the new IP subnet to the MT. Moreover, specific mobility protocols relying on mobility agent, like Mobile IP [5], may be required to maintain on-going sessions during handoff.

The two following sections will focus on the link level protocols allowing a MT to automatically change AP for two standardised WLAN systems, namely IEEE802.11 and
HiperLAN type 2. Afterwards will be exposed the mobility management based on IP protocols, that are not specific to WLAN but that are required in order to offer a complete mobility solution for IP access networks based on WLAN.

2.1 Mobility management with IEEE 802.11

The IEEE 802.11 wireless LAN standard [2] only specifies ISO layers 1 and 2, i.e. the air interface only, and the IEEE 802.11 roaming support corresponds to a link-specific mobility protocol. The change of AP by a MT is performed as follows:

- A MT, that is associated to one AP and only one at a given time, periodically checks the radio signal level with its current AP as well as with other authorised APs characterised by the same identifier SSID (Service Set Identifier). The MT learns about these neighbouring APs by using two different scanning modes: in the passive scanning mode, the MT listens to each beacon channel transmitted by the different APs; in that way, it can identify the APs corresponding to the desired SSID; in the active mode, the MT has to send a specific probe message indicating the desired SSID.

- When the MT moves, it detects that the current radio signal weakens and it initiates a re-association procedure (link layer procedure) with the AP that provides the best radio signal level. The dynamic procedure for association and re-association allows the mobility of one MT from one radio cell to another one by moving the current association from one AP to another one, without forcing a new authentication phase between the MT and the new selected AP.

- The session continuity is then guaranteed, as far as the involved APs belong to the same IP subnet. The Ethernet switch or bridge that links both APs has just to learn about the new localisation of the MT and is automatically informed through broadcast information like ARP or Neighbour Discovery [1] messages. If the MT moves from one AP to another one connected to a different IP subnet (e.g. AP connected to another interface of the router), the MT has to acquire a new IP address and the handover between radio cells has to be managed with a mobility protocol like Mobile IP to offer a session continuity.

Moreover the standard is completed with the IEEE 802.11f under development [3], which specifies an Inter-Access Point Protocol (IAPP) and whose main purpose is to allow interoperability between multi-vendor access points. Regarding the mobility support, this protocol supplies the mapping between the layer 2 identifier and the IP address of the APs and the possibility to transfer context information about MTs between APs. The transfer of context information, like security parameters, aims at allowing faster re-association procedures after handover.

2.2 Mobility management with HiperLAN Type 2

The current HiperLAN 2 specifications [4] split an AP into two entities, the Access Point Transceiver (APT), which is the transceiver part of the AP and the Access Point Controller (APC), which represents the control functionality of the AP. Moreover, concerning the addressing issue within an HiperLAN 2 access network, a MT is characterised by a MAC identifier (MAC ID), that is unique in the scope of the AP the MT is attached to; an AP is assigned an AP identifier (AP ID), that is unique for a given network belonging to one operator. Then three types of handover are optionally supported by AP and MT: sector handover when the MT only changes the antenna sector of its AP; the radio handover when the MT changes its radio cell moving from one APT to another one belonging to the same AP; finally network handover when the MT changes its APC. The two first cases only involve one AP whereas the latter one implies the change of AP. Therefore in the following, focus is put
on the network handover case for HiperLAN 2, that may require a re-negotiation of parameters of the on-going user connections at the level of layer 2.

The network handover procedure takes place as follows:

- The MT permanently measures the radio level with its current AP and with neighbouring APs; during this measurement phase, the MT is in the Absence state.
- When it detects that the current radio level weakens, it tries to associate with the new selected AP, optionally informing the old AP of its imminent handover. With the Association procedure, the MT provides the new AP with the context of the active connections at the old AP (e.g. the old MAC ID, old AP ID). Then the new AP may reject the association request.
- After the Association procedure, the MT triggers the Link Capability procedure in order to negotiate the parameters of the association with the new AP.
- Depending on the capability of the fixed network to interconnect both involved APs, the new AP may exchange information with the old AP about the context of the MT for security and authentication purposes. Otherwise it has to repeat the complete encryption and authentication procedures.
- Finally, before completing the handover procedure, the new AP may either agree to keep the parameters of the connections established at the old AP or ask to set up completely new radio connections.

In the previous section, it has been assumed that an handover was triggered by the MT. However in the HiperLAN 2 standard, there is also an option that allows an AP to initiate an handover.

In conclusion, the existing wireless LAN standards may support in a similar way roaming between APs, as far as the APs belong to the same wireless network (e.g. same SSID for IEEE802.11 networks). The coverage of one access network based on IEEE802.11 or HiperLAN Type 2 may be extended by installing several adjacent APs. However these technologies alone do not allow real mobility services since they do not specify any handover mechanism neither to transfer information relative to a MT between APs nor to buffer user data until the MT is associated to its new AP. Therefore ongoing connections of a MT performing handover will suffer from delay and packet losses due to the time-consuming re-association (possibly requiring re-negotiation of the connections). The communications are even interrupted when the MT handoffs between two APs belonging to different IP subnets. In that case, layer 3 mechanisms are required to maintain layer 3 connectivity and to update the routing path towards the MT. The following sections focus on these mechanisms.

3 Mobility management based on IP protocols

These last years, the Internet network originally based on the interconnection of fixed LAN became very popular and made IP protocol essential for the development of telecommunication networks. At the same time, the miniaturization of data-processing equipment like portable computers and the development of wireless networks and services increased the possibility for mobility. An IP mobile terminal, connected to the Internet, is localised according to its attachment point with its IP address; if it moves to a new attachment point, it must imperatively:

- change its IP address: it is for example the solution adopted by the dynamic address allocation scheme allowing the user to be nomad over several sites; this solution however requires to stop all IP transfers in progress; it thus allows to be nomad but does not provide service continuity during mobility
or inform all the routers of the Internet that its IP address locates a new attachment point, it is a non-scalable, i.e. unrealistic, solution.

Mobile IP [5] is the standard protocol for the support of macro-mobility i.e. mobility between IP networks. It allows a transparent routing of IP packets destined for MT in the Internet and ensures consequently service continuity for communications in progress. However, if Mobile IP is used to manage micro-mobility, i.e. mobility restricted to the same access network, it results in introducing delays in the diffusion of the new location and generates significant control traffic in the Internet core. These last years, several IP micro-mobility protocols have been developed to work out these problems. The current section consists of four parts. The first one briefly describes Mobile IP version 4 with a short comparison with Mobile IP version 6. The second part introduces other solutions proposed in order to provide mobility support to users. The third part presents two different categories of IP micro-mobility protocols [6] and their limits. The last part focuses on the IP micro-mobility protocol enhancements currently proposed within the IETF.

3.1 Mobile IP

Mobile IP protocol is the current standard for supporting macro-mobility in IP networks i.e. host mobility across IP domains while maintaining transport level connections. It is transparent for applications and transport protocols, which work equally with fixed or mobile hosts. It can be scaled to provide mobility across the mobile Internet. It allows nodes using Mobile IP to interoperate with nodes using the standard IP protocol. There are two versions of Mobile IP: Mobile IPv4 and Mobile IPv6. Each one addresses a particular version of IP.

3.1.1 Mobile IPv4

Mobile IPv4 protocol [7] defines three functional entities: the MT, the home agent (HA) and the foreign agent (FA), see Figure 2. The MT is configured with a permanent IP address belonging to its home network. It is called the home address. All packets sent to the MT are addressed to its home address. The home agent is a router in the home network of the MT. It is aware of the MT current location, when the MT is located outside its home network. The foreign agent is a router in the visited network. The MT uses it to obtain a new temporary address and generally to register it with the home agent. The home and the foreign agents are called mobility agents.

Figure 2: Functional entities according to Mobile IPv4

Figure 2 : Entités fonctionnelles mises en œuvre dans Mobile IPv4
Mobile IP performs as follows: when the MT is connected to a network, it listens to mobility agent advertisements broadcast by mobility agents. If the network prefix changes, the MT detects that it has moved. Then the MT is aware that it is located in a visited network and tries to acquire a new temporary address. This new address can either be obtained by an auto-configuration mechanism like DHCP [8] or be the actual foreign agent address. The former is called Co-located Care-of Address (CCOA) and the latter is called Care-of address (COA). If CCOA is acquired, the MT registers this new temporary address with the home agent by exchanging registration requests and responses using CCOA as the source address. If COA is acquired, the MT cannot register itself using this address. In this case, the foreign agent will perform the registration with the home agent. Once registration is completed, the home agent intercepts packets sent to the MT and uses IP-in-IP encapsulation [30] to tunnel them to the new temporary address. The home agent must also answer to the ARP requests destined for the MT hardware address with its own hardware address in order to intercept packets generated inside the home network [31]. If the MT uses a COA, the corresponding foreign agent decapsulates the packets and delivers them to the MT. Otherwise, the MT decapsulates its packets itself, as it is directly reachable using the CCOA.

To maintain the registration, the MT has to periodically renew it. When the MT returns to its home network, it has to remove its current registration. Therefore, the home agent will stop intercepting traffic destined to the MT.

In the basic mode, the home agent intercepts all packets addressed to the MT and tunnels them to the visited network. Consequently, a “triangular routing” effect is produced (Figure 3): all packets must first pass through the home agent even if the current access router is in the same network as the correspondent node. An extension of Mobile IP, known as Route Optimisation [11] (Figure 4), has been proposed to overcome this problem: it allows data packets to be routed directly from the correspondent node to the MT using a binding cache in the correspondent node that keeps track of the current temporary address. These binding caches are created and updated by Binding Update messages sent by the home agent or the MT in response to MT warnings or correspondent node requests.

![Figure 3: Triangular routing](image)

*Figure 3: Routage triangulaire*
Furthermore, with Mobile IP, the MT has to use its home address as a source IP address in order to maintain the current communications, i.e. it sends its packets through a router on the visited network, and assumes that routing is independent of the source address. Nevertheless, due to security concerns, the use of routers that perform ingress filtering [9] imposes on the MT the use of a source IP address that belongs to the network the MT is attached to. Consequently, an extension of Mobile IP, known as Reverse Tunnelling (see Figure 5, [10]), has been proposed to establish a topologically correct reverse tunnel from the care-of address, i.e. either the MT or the foreign agent depending on the temporary address of the MT, to the home agent. Sent packets are then decapsulated by the home agent and delivered to correspondent nodes with the home address as IP source address.

3.1.2 Comparison of Mobile IPv4 with Mobile IPv6

The design of Mobile IP support in IPv6 [12] is based on the experiences gained from the development of Mobile IP support in IPv4, and the opportunities provided by the new features of IP version 6 such as an increased number of available IP addresses and additional automatic IP auto-configuration features. Firstly, in Mobile IPv6, the route optimisation process is integrated in the protocol. In fact, both the route optimisation and the registration procedure with the home agent are done by new defined Binding Updates destination options. Furthermore, Mobile IPv6 and IPv6 itself allow MT and Mobile IP to coexist efficiently with routers performing ingress filtering, as the MT uses its temporary address as the source address. The home address of the MT is notified in a Home Address destination option of the IP packet. Also the use of IPv6 destination options that carry optional information only addressed to the destination node, allows all Mobile IPv6 control traffic to be piggybacked on any existing IPv6 packet, whereas in Mobile IPv4 and its Route Optimisation extensions, separate UDP packets are required for each control message.
Finally, in Mobile IPv6, there is no longer any need to deploy foreign agents. MT makes use of the enhanced features of IPv6 to operate in any location away from the home network without any special support required from these local routers.

3.2 Mobility support using other IP protocols

The support of Mobile IP is not an easy thing to do; it requires the deployment of a protocol which leads to additional delays and signalling overhead to communications. It also supposes an unwieldy security infrastructure. The standardisation process is constantly delayed as well. Then considering that only few long-lived TCP connections really need a transparent mobility management at the network level, it is suggested to use different kinds of mobility management depending on the real needs of the applications. Only important long-lived TCP connections will use Mobile IP for macro-mobility. Short-lived TCP connections, such as those initiated by a web browser, will not use any mobility support as it is sufficient to reinitiate the broken connections with the new IP address when the mobile terminal has moved. Other IP transport protocols are not facing with the same problems as TCP. In SCTP [32], it is possible to provide several IP addresses to identify a SCTP connection, which allows maintaining this connection while moving. At this time, it is not possible to change source or destination addresses during a connection but a draft proposes a dynamic address reconfiguration for SCTP [33]. For all applications using UDP, as there is no transport connection, there is no problem at transport level for maintaining a connection. Indeed, it is always necessary to manage mobility: detect the movement, get a new IP address, provide the new address to transport levels. But arguing that the application level is the better place to know which flows need to be transparently maintained while moving, it is suggested to manage the mobility at the application level. An example of this approach is the suggestion to enhance SIP to provide mobility support for multimedia flows.

3.2.1 Very brief description of SIP

The SIP architecture [34] already allows users to be reachable when they are out of their home network. When a correspondent attempts to open a SIP session (multicast, VoIP, video streaming) with another SIP user, it sends an INVITE message to the public address of the user. The format of this address is very similar to an email address: bonnin@enst-bretagne.fr which means the user "bonnin" in the domain or on the machine "enst-bretagne.fr". The correspondent uses a standard DNS to reach a SIP server that could be a redirect server or a proxy server. This SIP server plays the role of a Home Agent in Mobile IP. If it is a proxy server, it relays the INVITE message to the machine where the user is logged on. If it is a redirect server, it simply sends the new location to the correspondent in order to allow it to reach directly the user. As in the Mobile IP case, the mobile user is supposed to register its new location to its home SIP server each time it moves. This kind of mobility is called personal mobility since a user can change the terminal instead of simply moving with its terminal. All SIP messages are in ASCII format in order to keep the protocol and the treatment fast and simple.

3.2.2 Mobility support using SIP
With this way of working, all sessions are broken when a user or a terminal moves to a new location. A proposition [35] has been made to add to the SIP architecture the support of terminal mobility while sessions are active. It uses a method very similar to mobile IP route optimisation. When the user having SIP sessions active moves, it sends new INVITE messages (with the same session ID) to all correspondents for all active SIP sessions (Figure 6). This INVITE message contains the new location dependent IP address. It also sends a register message to its SIP server as previously described. In this way, all correspondents get the new IP address and further messages will use this new address as destination address. As we have already stated, this is not sufficient for TCP connections but it could be helpful for UDP or SCTP sessions. This proposition does not introduce new security considerations since security has been already taken into account in the SIP protocol and architecture. Using SIP with mobility support does not solve all mobility concerns. For instance, nothing is done to avoid packet loss while moving and nobody is able to relay messages arriving at the old location. This is the job of micro-mobility mechanisms dedicated to provide seamless handover inside a limited micro-mobility domain.

![Figure 6: Mobility support using SIP](image)

**Figure 6 : Support de la mobilité basé sur SIP**

### 3.3 IP micro-mobility protocols

Micro-mobility, i.e. mobility restricted to the same access network or limited to the same site, is characterized by local, frequent and fast displacements. If Mobile IP is used for the management of this movement type, the MT must emit registration request and binding updates at each displacement, even local, towards its home agent and its correspondent nodes, that may be located very far from its visited network. The IP network core is then loaded by update messages, a delay is generated for the diffusion of the new location and the communication suffers from long interruptions with significant packets losses. A micro-mobility protocol is necessary to mitigate these disadvantages. These last years, several solutions were developed in order to locally manage mobility of terminals and consequently to limit the diffusion of update messages within the network and to limit as well the delay required to take into account these messages. These IP protocols may be divided into two categories, those based on a hierarchy of mobility agents (or proxy agents) and those using localized enhanced-routing.

#### 3.3.1 Proxy agents architecture protocols
These protocols extend the idea of Mobile IP by using a hierarchy of mobility agents. A MT is registered with its local agent at the lowest level of the hierarchy; then the local agent is registered with the higher agent in the hierarchy and the registrations go up the hierarchy until reaching the home agent. When the MT moves, it emits only one registration request located in a domain. The packets that are destined for this MT, follow then the hierarchy of the mobility agents and are delivered from an agent to another via a tunnel. The most known protocols are Mobile IPv4 Regional Registration [13] and Hierarchical MIPv6 mobility management [14].

3.3.2 Localized enhanced-routing protocols

These protocols define the network architecture in two levels: the higher level is standard Internet network implementing Mobile IP and the lower level is composed of domains able to manage micro-mobility. The router connecting the two levels masks local movements compared to the rest of Internet network. Inside a domain, a specific routing protocol is introduced. The packets emitted by the MT update the routing entries in the intermediate nodes. Then a routing table entry maps the MT address with the address of the neighbouring node having transmitted the last packet responsible for the entry update. The chain of these correspondences represents the path traversed by the packets destined for the MT. The most known protocols are Cellular IP [15] and HAWAII [16].

3.3.3 Limits of micro-mobility protocols

These protocols of micro-mobility offer powerful solutions for the support of local mobility. They present the advantage to solve the principal limits of Mobile IP without complicating the mechanism of mobility management. Moreover, improvements are still possible especially with regard to the handover execution decision phase, the movement detection procedure and the handover execution time.

Indeed, in these solutions, the decision-making phase of handover execution is excluded from the protocol specifications. A MT undergoes simply the handover at the radio level: it changes access point each time the radio connection with its current access point is degraded. When the MT sticks to its new access point, it must wait for the reception of router advertisements in order to detect its movement and start the location update. Thus there is a considerable delay between the physical movement of the MT and its detection at the IP layer by the node itself, which obviously increases the handover execution time.

3.4 IP mobility protocol enhancements proposed for handover latency optimisation

Since the previously presented protocols still generate a delay due to the detection of movement at the IP layer in the MT, other solutions are currently discussed within the IETF. For example, Fast Handover for MIP [17] aims at optimising the handover latency by allowing the MT to acquire its new IP address before re-associating to the new IP subnet. This new address corresponds to the target attachment point and is then immediately valid once the MT is associated with this new attachment point at layer 2. Compared to the MIP standard, this FHMIP solution establishes new messages between the two access routers involved in the handover. Consequently, it allows the network to anticipate the new routing path towards the MT.

Alternatively, whenever the new IP address acquisition mechanism is not completed before the handover, the proposal [17] also suggests mechanisms to enable the MT to temporarily use its old IP address, at the new attachment point, corresponding to the old attachment point until it definitively acquires its new IP address, by installing a temporary tunnel between both attachment points. In an extreme case, the MT is able to successively move from one IP subnet to another one without changing its IP address, thanks to a temporary chain of tunnels.
between the different visited subnets. This alternative solution, called BETH (Bi-directional Edge Tunnel Handover), is mainly specified for MTs moving rapidly and involved in latency-sensitive real time communications.

Most of these enhancements proposed for IP mobility protocols rely on the existence of layer 2 (L2) triggers that are globally notification from layer 2. These L2 triggers allow to exchange information between layer 2 and IP layer within access network nodes (typically access router and MT) about imminent or recent events occurred at the radio link level. By the way, current works within the IETF show importance of various L2 handover triggers [18] and the need for close L2/L3 communications [19] in order to optimise handoff. However the way to implement these L2 triggers, that may be complex, is not clearly specified. Also these methods based on L2 triggers assume that the selection of the target attachment point is performed at layer 2.

Moreover, the IP mobility protocol proposals are based on the use of candidate access router discovery algorithms, which have been recently discussed within the IETF (for instance [20], [21], [22]), to select the best target attachment point for handover. These specific algorithms have to be implemented within MTs or access routers according to whether the handover is supposed to be initiated either by MTs ([20], [21]) or possibly by the network ([22], [23]). However these very recent algorithms still have to be evaluated. Also they require the support of specific protocols between access routers (e.g. specific information exchange between access routers, like mapping between link-layer identifier and IP address or capabilities of a given access router), which may be complex and time-consuming. Finally, these proposals do not precisely detail how handovers are initiated (either by the MT or the network) or how the list of candidate attachment points is established.

In conclusion, among the mentioned proposals for handover management, two cases may be distinguished:

- either the selection of the target radio cell during handover is performed by the MT itself taking into account radio measurement only, as it is also suggested in [24]. This solution may require heavy calculation within the MT and prevents the network from managing handover in order, for example, to perform load balancing between different IP subnets. Also, a MT may be redirected towards a radio cell that is already overloaded;
- or the handover may be initiated and controlled by the network, as the current access router is assumed to be able to select the optimal target attachment point for handover, taking into account more information, like radio cell resources, MT’s profile… This solution relies on specific access routers within the access network.

In the following section, focus is on our proposal, which is a network-controlled handover management approach. Indeed, it is considered that in case of large scale deployments, a network operator requires that the overall network and radio resources are controlled by a centralised network entity that is able to take into account various information like traffic load repartition as well as QoS parameters and user’s profile. Therefore, the selection of the target attachment point must not be restricted to layer 2 but rather be managed at the network layer.

4 Network-controlled approach for IP mobility management

Our approach [25] relies on the implementation within the IP access network of a new network entity called mobility manager (MM), see Figure 7, that is able to initiate and monitor MT handovers according to various parameters either static or dynamic (e.g. network configuration, current load of access routers, MT’s priority, user’s rights…). These parameters are stored and possibly periodically updated in a database, which is associated with the MM. A MT monitors the link quality with its current access point. When the MT detects that the
quality is falling below a given threshold, T1, configurable by the network, it collects quality measurements with neighbouring access points and reports them to the MM. In response, the MM selects a target access point to which the MT has to handoff by taking into account parameters stored in its database. In parallel, it duplicates packets destined for the MT and redirects them towards the future MT location.

This way, the MM is allowed to control the whole access network resource while anticipating MT handovers, which reduces handover execution time and packet loss. The MM can possibly completely eliminate packet loss by buffering data packets addressed to a MT, until the MT executes its handover and is reachable at its new attachment point. The choice to buffer packets (for example for TCP flows) or not (for example for UDP flows) may be justified by the type of the communication flow and its sensitivity to delay and packet loss. Besides, when a MT enters an access network managed by a MM, it is provided with a list of suitable neighbouring attachment points it is allowed to attach to. This list, which results from a pre-selection made by the MM, is specific to the MT and to its current attachment point. Consequently, the processing time of the radio measurements performed by the MT is limited as well as the overhead over the radio link due to the measurement report sent to the MM. Moreover, handover triggers within a MT simply consist in detecting the time when the current signal quality is falling below a certain threshold contrary to previously presented mobility management protocols that rely on specific L2 triggers. Indeed, the drivers of the usual wireless LAN cards directly provide signal quality indicators, which can be extracted and analysed. Then, the procedure of handover triggers is easily achieved.

Finally the suggested solution does not require any specific functionality within the access routers.

Below an example of simple algorithm of target access point selection allowing load balancing and priority management is given:
1. The MM maintains in its database information about the current load of the different access points belonging to its access network and about the relative priorities of the MTs covered by the access network.

2. The MM receives a handover request from a MT (MT₁), consisting of a radio measurement report related to the neighbouring access points of MT₁ sent by MT₁ to the MM.

3. The MM identifies MT₁ and analyses the radio measurements.
   3.1 The MM selects for MT₁ the access point offering the best radio link quality.
   3.2 If the access point is overloaded, the MM checks whether another MT (MT₂) with a priority lower than that of MT₁ is currently served by this access point.
      3.2.1 If so, the MM validates its selection and asks MT₂ to move to another access point.
      3.2.2 Else it selects the next best access point and reiterates the procedure from 3.2.

4. The MM returns its selection within its handover response sent to MT₁.

The next section will describe an implementation of IPv6 hierarchical mobility controlled by the network and will assess the enhancements regarding the handover delay.

5 NC-HMIPv6, a network-controlled mobility management implementation

The NC-HMIPv6 (Network-Controlled HMIPv6) protocol presented in [25] is a new IP micro-mobility protocol extended from HMIPv6 [26], [14]; HMIPv6 is the IP micro mobility protocol currently implemented and investigated within the France Telecom R&D test platform. Actually, HMIPv6 is one of the only IP micro mobility protocols considered in the IETF Mobile IP working group, and, moreover, its interoperability with Mobile IP is clearly specified [14], which makes it particularly meaningful. NC-HMIPv6 includes, in the mobility management protocol, the handover decision-making phase performed by the network. In [25], a short comparison between the performances of NC-HMIPv6 and HMIPv6 is given. This paper completes this first analysis with detailed experimental results.

5.1 Description of the test platform

Within France Telecom R&D laboratories, a test platform has been set-up, dealing with IP mobility issues. The platform is located in the Rennes laboratory. An IPv6 infrastructure has been deployed and provides native IPv6 interconnection.

The hardware infrastructure relies on:
- Desktops PCs for Home Agents (HA), local Mobility Agents (MA) and Access Routers (AR);
- Laptops for Mobile Terminals;
- 802.11b WLAN PC cards and Access Points (Cisco and/or Orinoco products), used in infrastructure mode;
- Ethernet hubs and switch;
- Routers for sites interconnection.
The test bed is configured as detailed in Figure 8. Two micro mobility domains have been set up. Access Points have been attached to the desktop routers (Access Routers, Home Agent, Mobility Agents).

![Diagram](image)

**Figure 8: France Telecom R&D test platform**

**Figure 8 : Plate-forme d'expérimentation de France Télécom R&D**

Essentially, the platform allows testing and demonstrating a number of mobility schemes, such as basic MIPv6 global mobility, intra and inter HMIPv6 domain mobility, mobility between a Home Network and an HMIPv6 domain. The HMIPv6 protocol stack in use on the platform has been provided by INRIA [26] and can be downloaded on [27].

One of the main objectives of the trial was to evaluate the relative performances of MIP, HMIP and NC-HMIP, particularly as far as real time multimedia communication is concerned. That is why, besides the classical legacy applications (ftp, web browser), a MPEG video server is implemented on a correspondent node (CN) within the network. It is able to serve video UDP datagrams. Client applications are installed on the MT. The Client initiates the communication, which is supported by two sockets, one for the control, one for the video data; in that way, the MT is able to remotely select the program, as well as to request at any time a start and stop in the video streaming.

A significant number of mobility scenarios can be tested on the platform, depending on the relative position of the correspondent node (CN), the old AP and the new AP. In this paper, we will concentrate on results that address micro mobility issues: it is the situation when the MT changes AP, and AR interface (i.e. change in the IP address), within the same micro mobility domain, while having an on-going communication with a CN located outside this domain.
5.2 Description of the experimentation

In the test platform, several independent parameters (relative locations of the CN, MT and HA, as well as type of communication etc…) can be chosen, which leads to a significant number of possible scenarios. A list of relevant configurations was set-up. This set of configurations will allow meeting the following objectives:

5.2.1 Demonstration of a handover using MIPv6

A Home Agent is used to register location of Mobile Terminals roaming in a visited network. The micro mobility protocol is not enabled on the platform. The Mobile Terminal network performs a layer 2 and layer 3 handovers, and a set of measurements is carried out.

5.2.2 Demonstration of the improvement brought by a micro mobility protocol (HMIPv6)

In the same configuration as when experimenting MIPv6, the Mobile Terminal experiences the same type of handovers, but this time, the visited access network enables the use of a micro mobility protocol (HMIPv6).

5.2.3 Demonstration of the enhancements brought by Network Controlled approach

To demonstrate the expected improvements, the same types of handover are tested with both micro mobility protocols, HMIPv6 on the one hand, and NC-HMIPv6 on the other hand. The measurements obtained on both configurations are compared. During the 3 types of demonstrations mentioned above, the handover delays were measured with UDP flows (Video Server) thanks to proprietary software implemented on the Mobile Terminal.

5.3 Experimental results

The different mobility scenarios are highlighted in Figure 9 and Figure 10.

Scenario 1 corresponds to a Layer 2 handover.
Scenario 2 corresponds to a Layer 3 handover on 2 different interfaces of the same access router.
Scenario 3 corresponds to a Layer 3 handover, on 2 different access routers in the same HMIP domain.
Scenario 4 corresponds to a Layer 3 handover, on 2 different access routers in different HMIP domains.
Scenario 5 corresponds to a Layer 3 handover, from a visited network to the home network.
Scenario 6 corresponds to a Layer 3 handover, from the home network to a visited network.

The difference between scenarios 3 and 4 is achieved by the configuration of the micro mobility domain border. In scenario 4, 2 domains were configured so that the old access router and the new access router belong to 2 different domains, while in scenario 3, only one domain including both access routers was configured.

Figure 9 and Figure 10 shows the 2 different configurations and the corresponding scenarios.
5.3.1 Measurements using an UDP application
One of the main objectives of our work on WLAN is to provide real time multimedia services to the user. This is the reason why we concentrated a significant part of the trial on real time communication.

Accordingly, this section highlights the trials that were carried out on the basis of an UDP video application, which includes an application server installed on a correspondent node. The correspondent node is always located outside the micro mobility domains, to avoid direct communication that would take place in that situation, between the mobile and correspondent nodes.

The overview results corresponding to Scenarios 1 to 3, (where micro mobility is helpful), are displayed hereafter in Table 1. Actually these 3 scenarios correspond to MT displacements inside a micro mobility domain and therefore constitute the most interesting cases of our study.

The UDP video server was configured to deliver a fixed rate video stream using 33 kbit/s, with a packet size of 1024 bits. A random delay of 1 to 4s has been set-up between two successive Router Advertisement messages. The correspondent node is located locally, but outside the micro mobility domain. Several measurements have been done for each scenario, which explains why the standard deviation is displayed as well.

<table>
<thead>
<tr>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(No change in IP address)</td>
<td>(Handover between 2 interfaces of the same AR)</td>
<td>(Handover between 2 AR within the same Access Network)</td>
</tr>
<tr>
<td>Handover delay</td>
<td>St. Deviation</td>
<td>Handover delay</td>
</tr>
<tr>
<td>MobileIPv6</td>
<td>0,316 s</td>
<td>27 ms</td>
</tr>
<tr>
<td>HMIPv6</td>
<td>0,282 s</td>
<td>59 ms</td>
</tr>
<tr>
<td>NC-HMIPv6</td>
<td>0,313 s</td>
<td>63 ms</td>
</tr>
</tbody>
</table>

Table 1: Handover execution times

5.3.2 Measurements using a TCP application

The wireless access network should as well support TCP type applications. A typical TCP type application is the File Transfer Protocol (ftp) application.

We experimented the different scenarios with the 3 different mobility protocols. A FTP server was installed on a correspondent node located outside the micro mobility domain.

For each handover situation (corresponding to the different protocols and different movement types), the File transfer from the correspondent node to the Mobile Terminal had always been completed. The transfers had been suspended for a few moments corresponding to the handover execution, then the transfers resumed and completed in a successful way.

5.3.3 Main results observed

- Each movement type is functionally working, i.e. the Mobile Terminal gets a new IP address, and after some time, is always able to receive IP packets from its correspondent node, with which it has started a real time communication. According to the scenario and the mobility protocol, the communication is suspended during a moment, but is always able to be resumed.

- Scenario 1, which corresponds to a Layer 2 handover, since both old and new access points are connected to the same interface of the Access Router, displays the same performance range for all 3 protocols MIP, HMIP and NC-HMIP; this confirms what we expected.
- Scenarios 2 and 3 underline the difference between the mobility protocols. The examination of the results (Table 1) confirms our expectations.
  o With MobileIPv6 and HMIPv6, the handover delay has the same range of value; there are no significant differences between both protocols, mainly because the communication with the Home Agent was very fast. HMIPv6 however saves some additional message exchanges with the Home Agent, which leads to fewer loads in the IP core network. However, in a real size network, with a Home Agent located far away from the MT, the results could be different however, and it can be expected that, in bad IP Core network conditions, and with a great number of MT, MobileIPv6 would lead to a longer handover delay, while HMIPv6 handover delays would remain the same, since the mobility is locally handled.
  
  o With NC-HMIPv6, the Mobility Anchor Point acts as a proxy, as in HMIPv6, but as the handover is planned, the Mobile Terminal as well as the Mobility Agent knows the future IP address before the handover is really performed, and therefore, the packets can be redirected immediately to the new address. The handover delays (and hence the packet losses) are much lower than with HMIPv6, and fall down to around 300ms, which is in fact in the same value range as with a simple layer-2 handover.

- When using the video client application, we experienced that the use of a micro mobility protocol enhanced with a planned handover functionality (as provided by NC-HMIPv6), is quite appropriate; actually the reception of the video stream is frozen during the handover duration time, which is barely noticeable with the typical 300ms value obtained with NC-HMIPv6. With the previous mobility protocols (Mobile IPv6 and HMIPv6), the service is not broken, but interrupted for a significantly longer period of time, which may not fulfil the expectation of a mobile user.

5.4 Simulation results using a TCP application

In order to complete the experimentation results relative to TCP applications, simulation studies based on the OPNET modeler version 7.0 have been performed. Both HMIPv6 and NC-HMIPv6 protocols have been implemented. All simulations are performed using the network topology shown in Figure 11.
The mobile node is configured with four links and is initially connected to its home sub-network through hub_1. That means that at $t_0$, only the link with hub_1 is activated. The mobile node then moves within a visited sub-network. The movement is simulated by deactivating the link between the mobile node and hub_1 and by activating the one between the mobile node and router_6.

The visited sub-network consists of three access routers and a NC-HMIPv6 mobility manager (or HMIPv6 local mobility agent). Once the mobile node is registered in the visited sub-network, it establishes a TCP connection with a correspondent node (server_1) for the transfer of a 500 Kbytes ftp file. The transfer begins at $t_0 + 60$ seconds. At $t_0 + 66$ seconds, the mobile node moves and is connected to router_7. The link of the mobile node with router_6 is then deactivated. Afterwards, the one with router_7 is activated. Between the old link failure and the new link activation, an interruption time of 0,275 second is enforced to simulate the radio handover.

TCP parameters used for the simulation are as follows: the maximum segment size is 1360 bytes; the receive buffer length is 13600 bytes (i.e the maximum number of segments sent is 10); the algorithms of Kann and of Nagle are enabled; the algorithm slow start is enabled with an initial count of one segment; the selective ACK is enabled and the maximum ACK delay is 0,2 second. Also, the network conditions are assumed ideal in order to suppress packet loss due to network congestion. This alternative was chosen in order to be able to simply concentrate on packet losses due to the mobile node mobility.

The simulation results are described in Figure 12. They show that the ftp file download response time in NC-HMIPv6 is better than the one in HMIPv6: 11,19 seconds for NC-HMIPv6 and 15,04 seconds for HMIPv6. The received segment sequence numbers curves explain that this improvement is the consequence of a shorter interruption in NC-HMIPv6. Moreover, we note that the HMIPv6 interruption causes a loss of segments since the
mechanism of slow start was activated again. It is not the case of NC-HMIPv6 where the segments are simply delayed. The analysis of simulation TCP traces show that:

- For HMIPv6, during the handover execution, 10 data segments arrive at the level of the mobility agent and are routed towards the old location. They are thus lost. Later on, the TCP server does not receive an acknowledgement and returns to the slow start phase.

- For NC-HMIPv6, the mobility manager is informed of the movement before the handover occurs. The 10 segments are routed towards the future location and arrive at the level of the new access router (router_7) where they are buffered. When the mobile node executes its handover, it emits a neighbor advertisement towards its new router. The segments are then retransmitted to it and the mobile node can consequently acknowledge them on time. For our simulation, the buffer of the new access router was long enough and no packet loss occurred.
6 Conclusion and perspectives

This paper focuses on the way to handle the mobility within IP radio access networks based on WLAN technologies. Moreover, it shows the interest, from an operator point of view as well as from a mobile user point of view, of a network-controlled handover management approach particularly in order to share the whole resource in large scale access networks. Then, a hierarchical IP mobility protocol supporting a handover management controlled by the network has been implemented. The experimental results have confirmed that the
network-controlled planned handover allows to reduce the handover latency to the radio (layer 2) handover execution time.

Particularly, the experimentations carried out on France Telecom platform confirmed the interest to implement an IP micro mobility protocol to improve the overall performance of a Mobile Access Network. Doing this way, the signalling load over the IP core network is considerably reduced, and in addition, the handover performances get better.

To further improve the micro mobility protocol performance, a "Network Controlled" extension has been implemented on the basis of an existing HMIPv6 implementation. This new protocol called NC-HMIPv6 includes a handover planning function, which further decreases the handover delay to values as low as 300ms. This allows the access network supporting real time application such as video conference or voice calls in a nearly seamless way.

The platform is open to further experimentations and evolutions since the "Network controlled" approach can apply to other "proxy based" micro mobility protocol. Particularly, in a next step, QoS management mechanisms could be investigated addressing the mobility manager as well as the mobile terminal. In addition to the handover planning feature, the Network controlled approach offers additional possibilities worthwhile for a network operator, such as the possibility to control and balance the load of the different access point, as well as implementing QoS features, by controlling and filtering the IP packets, according to their class of service and to the priority of the destination Mobile Terminal.

A detailed description of these mechanisms is outside the scope of this paper. However, examples are given below.

6.1 Dynamic configuration of the quality control threshold

The threshold $T_1$ of the handover execution procedure control can be dynamically configurable by the network administrator. Indeed, the mobility manager can continuously suggest to the mobile terminal an optimised quality control threshold, according in particular to the specificity and the current state of the cell managed by the current AP and to the type of the communication flow towards the mobile terminal. For example, if the mobile terminal is transferring video flow, which requires a high quality link, the mobility manager may suggest to the mobile terminal a rather high control threshold.

6.2 Filtering using service differentiation

Packets addressed to mobile terminals served by the same access point are firstly intercepted by the mobility manager; then they are delivered to these mobile terminals via the same access point. The mobility manager can anticipate the overload of the radio cell managed by the access point and filter the intercepted packets before delivering them. To this end, it can favour either packets addressed to mobile terminals with high priority or the most important packets according to IPv4 Type of Service field or IPv6 Class of Traffic field.

6.3 Admission control

Finally, the hierarchical mobility management controlled by the network is particularly suitable for telecommunication operators. Thanks to the mobility manager, operators can easily implement charging system and admission control at the IP level. Since the mobility manager always intercepts packets sent to mobile terminals, operator may for example impose on mobile terminals the use of reverse tunnelling via the mobility manager. Operators may then configure firewalls to accept only packets addressed from the mobility manager or towards mobility manager subnet. Thus, a mobile terminal, whose admission control fails, cannot communicate at the IP level.
References
[4] Technical Specification, "Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; Data Link Control (DLC) layer; Part 2: Radio Link Control (RLC) sublayer