Heterogeneous Network Handover
Seamless Mobility Provision

Erik P Hammarberg¹, Muhammad Hammad Saghir², Narjis Hachilif³, Vetriselvam Gopal⁴

¹Department of Computer Science and Engineering, ²Department of Signals and Systems
Chalmers University of Technology, 412 96 Göteborg, Sweden
{¹hammarbe, ²saghir, ³narjis, ⁴gopalv}@student.chalmers.se

Abstract — The fourth generation (4G) networks are based on the idea of IP-convergence, where a user can roam across different networks. This idea of heterogeneous networks is to provide higher quality of services (QoS) in terms of high data rates. But this roaming across the heterogeneous networks has an inherent problem of vertical handover. IETF has provided IP layer solutions, e.g. Mobile IPv6 (MIPv6), to this problem but these solutions have higher handover latency. To reduce this latency further and to ensure the seamless service provision IEEE has proposed a new standard, IEEE 802.21 Media Independent Handover (MIH), which handles the entire lower layer signalling. Thus, in this paper we are going to study the specifications of MIH, MIPv6 and Fast-MIPv6 (FMIPv6); and we will look how the mentioned protocols perform a vertical handover. We will also propose changes to the MIH and FMIPv6 that can further reduce the handover latency.

Keywords — 4G, Heterogeneous Networks, Vertical Handover, IP, Mobile IPv6, Fast_MIPv6, MIH, IP-centric, Latency.

I. INTRODUCTION

The wireless access technologies, such as CDMA2000, UMTS and WLAN, have tremendously evolved and deployed all over the world in the past few years. The user demands are better services including higher data rate, higher mobility and higher Quality of Service (QoS). The user demands cannot be fulfilled by each wireless access technology alone, e.g. WLAN does not provide high mobility but high data rates, while CDMA2000 and UMTS do not provide high data rates but high mobility. Therefore, the Future Generation Networks (FGNs) are integrating these kinds of multiple wireless access technologies into a common convergence network to fulfil a user’s demands. To support this service, there are certain crucial requirements, like very high data rates for ‘any-service’, multi-interface systems for ‘any-device’, integrated heterogeneous networks and high mobility for ‘any network’, and so on [1].

To provide continuous session continuity in a heterogeneous environment, a vertical handover is needed whenever a suitable network is found. Each of the technology developer define a framework for a horizontal handover but to perform a vertical handover IETF has provided a worldwide accepted solution known as Mobile IP (MIP). Over the period of time, MIP has evolved and many new-advanced version of it have been issued, MIPv6, FMIPv6, HMIPv6, Proxy-MIPv6 (PMIPv6), and implemented [2,3,4]. MIP has shown satisfactory performance but its handover latency is a challenge. To solve this issue, IEEE has provided a new standard IEEE 802.21 (MIH). This standard handles the entire lower layer signalling and MIP executes the handover. Thus in this way the handover latency is considerably reduced.

In this paper, our main focus is how to reduce the handover latency in order to provide the seamless mobility. The current standards, e.g. MIPv6 along with MIH which are discussed in this paper as well, provide the seamless mobility. But, in order to enhance the user’s experience of session continuity in a heterogeneous environment, we have presented a novel signalling scheme. This scheme presents the idea of improving the performance of FMIPv6, and its enhanced versions, by acquiring a reserved IP address early on, either by FMIPv6 itself or by MIH.

The rest of the report is organized as follow; section II presents the overview of existing IP layer solutions, MIPv6 and FMIPv6, section III presents the overview of MIH, section IV and V presents the proposals for improvement in FMIPv6 and MIH, respectively, and in section VI the whole paper is concluded.

II. IP-LAYER HANDOVER SOLUTION(S)

The MIP system’s basic principle is that a mobile node (MN) is always addressable with its home address, regardless of where it is connected to the internet. When connected to its home network, packets to the MN are routed in a conventional way between sender and receiver, but when the MN connects to the internet at a different subnet, MIPv6 will ensure that it is still accessible via the same address that it uses at its home network [5]. MIPv6 consists of 3 phases:

- Movement Detection (MD)
- Duplicate Address Detection (DAD)
- Registration Process (RP)

which causes the handover latency in MIPv6, numerically it is represented as [6]:

\[ t_{\text{MIPv6}} = t_{\text{MD}} + t_{\text{DAD}} + t_{\text{RP}} \]

Following is a description of MIPv6 and FMIPv6:
A. Mobile IPv6

The MIPv6 defines two different modes of operation, bidirectional tunnelling and route optimization. In both of these, the MN has a home agent (HA) resident in its home network. In order to be reachable at any time by a correspondent node (CN), the MN also has two addresses, a home address and a care-of address (CoA). The home address representing the fixed address of the MN, it identifies the MN regardless of its current location. The CoA is associated with the MN when it is away from its home network, and belongs to the network the MN is currently visiting [7]. When the MN is not present in its home network, it contacts HA to register its CoA to form a binding between the HA and the MN at its current CoA.

In the bidirectional-tunnelling mode, traffic to the mobile node arriving from a CN is routed to the HA and the home agent tunnels the traffic to the MN at its CoA and vice versa [5], as shown in figure 1.

In the route optimization mode, the MN is required to register its current binding at the CN. This enables the CN to use the direct route to the current CoA of the MN. By using a special type of IPv6 routing header, which includes the home address, the use of CoA and special routing options is hidden from the above layers [5].

Route optimization mode is more efficient in reducing the packet delay since it allows the use of a direct path. However, it does not provide location privacy with the communicating nodes as bidirectional tunnelling does [8].

Mechanisms involved in MIPv6 can introduce handover latency, packet losses and signalling load [7]. Several enhancements such as FMIPv6 and Hierarchical MIPv6 have been proposed to deal with these issues.

B. Fast Mobile IPv6

A major part of the latency experienced in MIPv6 can be due to initial setup and authentication [9]. A scheme for reducing this latency is described in [3] and is called Fast Handovers for MIPv6. The goal of this scheme is to reduce the latency, or delay, when the MN cannot send or receive packets. Reducing this period could both increase the performance of real time systems such as VoIP and reduce, and possibly avoid, packet loss during the handover [3].

The FMIPv6 introduces new mechanisms that enable the MN to search, authenticate and receive address information for a new network while still connected to the previous network. Thus the MN can register a binding and have the HA prepare tunnelling of packets even before the switch between networks takes place [10].

When a mobile node (MN) discovers a New Access Router (NAR), it sends a ‘Router Solicitation for Proxy’ (RtSolPr) message to the Previous Access Router (PAR). The PAR
responds with a ‘Proxy Router Advertisement’ (PrRtAdv) message containing information about the access point. With this information, the MN creates a new care-of address (CoA) and sends it through a ‘Fast Binding Update’ (FBU) message to the PAR. The purpose of this FBU is to authorize the previous router to bind old CoA and new CoA, so that arriving packets can be forwarded to the new location of the MN. The PAR then sends a ‘Handover Initiate’ (HI) message to the NAR, containing the MN’s old CoA and the proposed new one. The NAR responds with a ‘Handover Acknowledgment’ (HAck) message indicating that the proposed CoA is valid. The PAR can then respond to the FBU by sending a ‘Fast Binding Acknowledgment’ (FBAck) message to both the MN and the NAR. Depending on whether this FBAck has been received or not on the previous link, there are two modes of operation: predictive mode and reactive mode.

In the predictive mode, shown in figure 2, the MN knows that the PAR has received the FBU previously sent. The MN can then disconnect from the PAR in order to connect to the NAR. It sends a ‘Fast Neighbour Advertisement’ (FNA) message immediately afterwards, so that arriving packets can be forwarded back to the MN straight away. Finally, the MN sends Binding Updates (BU) to CN so that they can send packets directly to the new network [7].

In the reactive mode, shown in figure 3, the MN does not know if the PAR has successfully processed the FBU. Hence, it resends the FBU as soon as it connects to the NAR. This time, the FBU should be encapsulated in the FNA so that the NAR can tunnel packets immediately. Packets are then forwarded the same way as in the previous mode [3].

FMIPv6 minimizes the latency and packet losses observed in the MIPv6 handover. However, when saturation arises, FMIPv6 has shown a worse performance than MIPv6 [7].

III. IEEE 802.21: MEDIA INDEPENDENT HANDOVER (MIH)

Mobile IP, and its different versions, have an inherent problem of high handover latency due to extensive lower layer signalling, which affects the seamless mobility provision to a user. Moreover, in a heterogeneous environment, where different access technologies have different lower layer specifications, it’s not possible for a single IP layer to communicate with different lower layers at once. Thus, to overcome these issues, IEEE has developed a new standard called IEEE 802.21 or Media Independent Handover (MIH) services.

The figure 4 shows the general architecture of an MIH enabled entity [11]. The core entity in this architecture is MIH function, which acts as an abstraction layer and provides services to MIH user, the upper layers e.g. IP layer, through its SAPs. Through its SAPs, it provides three types of services;
- Media Independent Event Services (MIES): allows higher layer to receive all the lower layer events, either local or remote.
- Media Independent Command Services (MICS): allows the higher layers to configure and issue some tasks to lower layers.

Fig. 4 General Architecture of MIH

- Media Independent Information Services (MIIS): allows MIHF to collect all the static information of the surrounding networks.

These three services devise a framework of the flow of messages between the different entities involved in handover; figure 5 shows the framework of a mobile initiated handover [12]. The framework in figure 5 is a modified framework which is accepted by IEEE in Feb 2009. It consists of four phases; phase 1 starts when the MN gets attached to the network, it shares its handover profile describing the user requirements and priorities. Then the serving network’s MIHF gets all the static information about all the surrounding networks. In the network discovery phase, MN asks the
network to check the availability of resources at the new networks and select the best possible. In the network selection, the network asks for the reservation of resources at the new network. Once the resources are reserved, a link layer connection is established and serving network asks the MN to conduct the handover. The execution of handover is out of scope of MIH, so the MIPv6 conducts the handover at IP layer. Once the handover is performed, the MIH performs there source releasing at the previous network, in the MIH completion phase.

This framework allows the MIPv6 to perform better by handling the entire lower layer signaling, thus resulting in lower handover latency and better seamless service provision.

IV. NOVEL FMIPv6 SCHEME

The FMIPv6 scheme suggested in [3] does reduce the initial detection and setup latency. It does, however, nothing to reduce several other forms of latency and, in the case of the proposed new-CoA (nCoA) being rejected by the NAR, might be thrown back to the same performance as regular FMIPv6. To combat some of these shortcomings, we suggest two updates to the standard. These two improvements could be implemented independently of each other but greatest benefits would be to implement them together.

A. Network Address Reservation

The first suggested improvement to the protocol is a scheme for the MN to be able to reserve an IP-address at the NAR in advance. The possible workflow of this improvement is shown in figure 6. The suggestion is that the PAR, upon receiving the RtSolPr message, sends an Address Reservation Request (AdrResReq) message to the NAR. The NAR then subsequently responds with an Address Reservation Reply (AdrResRpl) message with an address for the MN to use when connecting. The protocol then proceeds like the original protocol with either sending a FBU message or switching network and then sending an FBU embedded in the Neighbour Advertisement (FNA) sent to the NAR.

The advantages with this scheme are twofold, where the first advantage is always present and the second is experienced when this improvement is used in conjunction with the second FMIPv6 improvement suggested in this paper. The first advantage is that the risk of the nCoA that the MN has constructed and which is used in the FBU will never be rejected. This of course provides greatest advantage to the situation where the FBU is sent via the NAR after the MN has already disconnected from the PAR, because if the address is rejected in this situation, the latency will fall back to the same as FMIPv6. This scheme, however, would also provide great advantage to fast moving nodes: it could be that the MN has the time to perform one FBU/FBAck with the PAR before the channel fades, but a second negotiation will not be possible over that channel.

B. Proactive Binding Updates

The tunnelling of packets during handover will always be a source for certain latency, since the packets will be routed to PAR, tunneled to NAR and then delivered over the link between NAR and MN. The second improvement we suggest attempts to decrease the amount of time over which this latency is incurred. In the current suggestion for FMIPv6, the BU is sent much as in the standard FMIPv6; this is done as the very last thing performed in the handover sequence. In error and that the MN is connected and IP-capable at the new network when the packets are routed here. If low latency is the goal of the scheme however, this approach could be viewed as too conservative as packets will be tunneled for a long time after the MN is IP-capable on the new network.

The improvement suggested to combat this issue is to send the BU to the CN as soon as the MN has confirmation of which IP-address will be used and has initiated handover. In the case of a non-improved version of FMIPv6 being in use, this information will be possible to send after receiving the FBAck from the PAR. If the above improvements regarding IP-address reservation were implemented in the protocol, the MN could possibly send the BU right after the FBU was sent.

Using this scheme would decrease the amount of packets that was routed to the PAR after the MN has disconnected and
thus having to be tunnelled to the NAR, possibly avoiding using the tunnel altogether. The main drawback with this scheme is that if BUs are handled quickly, the packets from the CN could arrive at the NAR before the MN is connected and thus have to be buffered (or dropped). This, however, does not have to be a negative, since the updates are sent as the MN is about to disconnect, the packets that arrive early at the NAR would, with a high probability, have been delivered at the PAR after the MN had disconnected and thus would be tunnelled and possibly being delivered later than had been buffered at the NAR.

V. NOVEL MIH SCHEME

A. Improved MIH Framework

The MIH framework, figure 5, has a plenty of room for improvements. In this paper, we focus on the seamless service provision by further reducing the layer-3 (L3) handover latency. Latency at L3, as discussed earlier, is due to 3 processes [---]:

\[ t_{\text{MIPv6}} = t_{MD} + t_{DAD} + t_{RP} \]

The latency due to DAD process is removed through optimistic-DAD (oDAD) and advanced-DAD (aDAD). So, in order to remove the latency due to MD, we propose two new MIH messages which will enable the MN to get a new and reserved IP address at the new network even before the MIPv6 conducts the handover, the figure 7 depicts this whole procedure.

So, when in the network selection phase the target network receive the message, from serving network, to reserve the resources for MN, it sends the MIH_IP_Address_Request message to its IP layer to get a new and reserved IP address for the MN. The IP layer gets a new IP through DHCP and sends it to MIHF via MIH_IP_Address_Response. The new network sends that IP address through modified MIH_N2N_HO_Commit Response message to the previous network and then is forwarded to MN. In this way MN gets a new IP address even before going through movement detection process of MIPv6 and MN can directly send it for binding updates to CN while in the mean time the two networks setup the tunnel. Thus, in this way we can eliminate the L3 latency due to MD process.

B. Performance Evaluation

From [6] and [12]:

\[ T_{\text{handover}} = 3D_{\text{MN,SAR}} + 6D_{\text{SAR,TAR}} + 2D_{\text{MN,TAR}} + 2D_{\text{SAR,IS}} + P_{\text{Net_decision}} + T_{L2} + T_{L3} \]

where,

\[ T_{L3} = t_{MD} + t_{DAD} + t_{RP} \]
is handover latency caused by IP layer, and $D_{X,Y}$ is the message delivery delay between X and Y. $P_{Net \_decision}$ is the process time of new network decision and $T_{L2}$ is the latency caused by link connection establishment.

Now, the latency equation for the proposed scheme,

$$T_{handover} = 3D_{MN\_SAR} + 6D_{SAR\_TAR} + 2D_{MN\_TAR} + 2D_{SAR\_IS} + P_{Net \_decision} + P_{IP \_address \_reservation} + T_{L2} + T_{L3}$$

where, $P_{IP \_address \_reservation}$ is the process time for IP reservation at the new network, and $T_{L3}$ is the latency caused by link connection establishment.

Normally, the L3 handover latency is measured about 2000 to 3000 ms while the layer-2 (L2) handover latency is measured about 100 to ms [13]. Considering those figures, we verify that proposed scheme considerably reduces the handover latency. Moreover, the process is same for a network initiated handover scheme.

VI. CONCLUSION

This paper discussed the potential problem of seamless mobility in heterogeneous networks. The existing L3 handover solutions, e.g. MIPv6, FMIPv6, along with MIH provide the seamless connectivity. But there is plenty of room for improvements in those solutions. In this paper, we have specified some novel messages, both in FMIPv6 and MIH, in order to improve the handover framework. The address reservation messages, both in FMIPv6 and MIH, enable the MN to get a reserved IP address before hand, thus reducing the latency at L3 due to movement detection and address detection. Moreover, by using the novel proactive binding update scheme, we can reduce the latency due to address registration both at NAR and CN. These schemes can be implemented one by one or together, but most advantage would be gained by implementing them in unison. The conclusion that can be drawn from the work performed in this paper is that issue of seamless handover between packet networks is still an open issue with room for improvement. The largest hurdle is perhaps not to develop the best technical solution but to gain consensus for an improved solution so that widespread usage of the scheme could begin.

REFERENCES

Review Question:

Q. What is seamless mobility provision in a heterogeneous environment; what are the issues and their reasons?

Ans. In a heterogeneous environment, where a user has access to different access technologies simultaneously, a user’s priority is the continuity of the ongoing session while on move. Thus, to ensure this session continuity in a fashion that the user doesn’t feel any interruption, e.g. service degradation, end of the session, etc. is referred to as seamless mobility provision.

But this mobility in heterogeneous environment induces a problem of handover between different access networks, called vertical handover. Each of the technology developer defines a framework for handover between same technologies, called horizontal handover, but vertical handover is quite complicated because different access technologies have different lower layer specifications. There are several standards that provide the session continuity, e.g. SCTP, SIP, Mobile IP, UMA, etc. but the difference of specifications causes increased lower layer signalling in those standards, in order to exchange the handover related information. This lower layer signalling causes the handover delay which affects the seamless mobility experience of a user.