Mobility Management in Next Generation Networks: Analysis of Handover in Micro and Macro Mobility Protocols

Shaima Qureshi, Ajaz Hussain Mir

1Department of Computer Science & Engineering, National Institute of Technology Srinagar, Kashmir, India
2Department of Electronics & Communication Engineering, National Institute of Technology Srinagar, Kashmir, India


Abstract: One of the main issues with Mobile Internet Protocol (MIP) is slow handovers. Mobility protocols are classified as micro or macro depending on the domain of the network. When macro mobility protocols are used in managing localized mobility requirements, they result in slow handovers and delays that result in loss of packets and make these protocols unsuitable for time sensitive applications. Micro mobility management protocols have been proposed to resolve this issue. Some of these localized protocols are more attractive as they keep mobility restricted to the network that removes the need of having mobility management support in their software stack. In other protocols hosts are involved in mobility management. In this paper we will review the concept of mobility and the mobility protocols available in IPv6. We will discuss proposed protocols aimed at macro and micro mobility management by grouping them according to their host or network based management approach. We will also compare their handover performance.

Keywords: Mobility; Micro – Mobility; Macro- Mobility; Mobile Internet Protocol version 6 (MIPv6); Handover latency; Local Mobility; Global Mobility; Host-Based Mobility Management; Network-based Mobility Management.

1. INTRODUCTION

With advances in technology, the number of mobile devices being used is increasing tremendously. Increasing use of mobile devices in the internet has generated the need for a new protocol (Mobile internet Protocol, MIP) as well as an expanded address space for identifying such devices. Mobility answers the requirement of seamless movement of these portable devices. Mobile Internet Protocol (MIP) was approved by the Internet Engineering Steering Group (IESG) in June 1996 and published as a Proposed Standard in November 1996. Mobile IP is the earliest solution to mobility management of IP network. The complete description of this protocol is given in IETF RFC2002. The need for new technologies and Quality of Service requirements and the shortage of IP addresses has evolved Mobile IP to Mobile Internet Protocol version 6. The technologies of General Packet Radio Service (GPRS), Wireless Local Area Network (WLAN) and CDMA2000 1X are boosting the development of Mobile IP, providing a platform for the implementation of Mobile IP services [1,2].

When a mobile node (MN) moves from one network to another, its IP address has to change to reflect the new network it has joined. While doing this transition current sessions and connections need to be maintained. For higher level protocols like Transport Control Protocol (TCP), any change in the IP address or port numbers is detrimental to the ongoing sessions as it breaks the continuity of the session and goes against the concept of Mobility in IP. A break in the connection should not happen within the mobility supported protocol. In mobility, when a mobile device shifts its link-layer point of attachment to the Internet, it should not change its original IP address, i.e. the home address of the MN. A MN must also be able to communicate with other nodes that do not implement these mobility functions [3,4]. When mobility in Internet Protocol (IP) was devised it was kept in mind that such devices are not connected by physical media and as such may have lower bandwidth and higher error rates. Also such devices are likely to be operated on batteries and as such minimal power should be consumed when mobile. The administrative message to be exchanged during the
process of mobility had to be minimized as well and their size had to be kept nominal to avoid burdening the network with exchange signals that might occur frequently [5]. Mobile Internet Protocol version 4 (MIPv4) introduced agents in the home and foreign networks that helped in transparent movement of the mobile devices between networks. The transfer is done in such a way that the higher transport layers receive the original IP address of the MN, retaining the connection channel [6]. Thus the session continues even if the network location changes for a MN.

Mobility Management protocols can be host based or network based. MIPv6 is a host based mobility protocol in which the MN plays a role in the mobility scenario. Host based mobility allows a MN to change its point of attachment to the network, without interrupting IP packet delivery to or from the node. The current location of all the MN’s in the network is maintained by Access Network Procedures. Host mobility is also known as terminal mobility [7]. Other host-based mobility protocols include HMIPv6 (Hierarchical Mobile IPv6), FMIPv6 (Fast MIPv6), F-HMIPv6 (Fast Handover for Hierarchical MIPv6). In a network based mobility management, the MN is freed from any mobility related activities. It allows an entire network to change its point of attachment to the Internet. Thus IP packet delivery is not affected as it is still reachable in the topology. Proxy MIPv6 (PMIPv6) is a Network based mobility management protocol and the only network based mobility management protocol standardized by IETF.

Apart from the mobility protocols being classified as host and network based, there are two further subcategories within host and network mobility. First category is the global mobility protocols also termed as the macro mobility protocols [7]. MIPv4 and MIPv6 are both macro mobility protocols. Macro Mobility Management Protocol is a mobility protocol that maintains session continuity when a MN moves from one network to another causing change in its network topology [8]. The global, end-to-end routing of packets is changed during mobility to maintain session continuity. Apart from MIPv6, there are some other macro mobility management protocols that the IETF is working on and as such future wireless networks might support more than one mobility protocols simultaneously. Host Identity Protocol (HIP) [9] and IKEv2 Mobility and Multihoming (MOBIKE) [10], are some macro mobility management protocols [8].

PMIPv6 (Proxy Based Mobile IPv6) is an example of a localized mobility management protocol also known as micro mobility protocol [5]. Mobile IP is not designed to support fast handoff in handoff-sensitive environments. It produces a lot of control traffic inside the local domain that increases handoff delay and the risk of packet loss. As such it is not suited for mobility between pico-cells and real-time application usage. There is significant signaling overhead, handover latency, and transient packet loss and are jointly known as fast handover issues. For each handover, signaling has to take place between mobile host and its home agent, which takes time and adds to the network load. The signaling load is proportional to the number of users and their level of mobility [11]. These generate the need for micro-mobility schemes that can satisfactorily handle localized movement without any support from wide-area protocols. A number of micro-mobility schemes like Cellular IP, Hawaii, Hierarchical Mobile IP, Intra Domain Mobility Management Protocol (IDMP), Edge Mobility Architecture etc have been proposed over the last couple of years, an overview of which can be found in [12]. While many of those proposals address the fast handover issues with a good degree of success but they lack flexibility and the capabilities for QoS and gradual deployment. [13]. A significant number of Internet Service Providers and network operators are migrating towards Multi Protocol Label Switching (MPLS) [14] as the transport option for IP services. MPLS provides notable benefits like QoS, Traffic Engineering (TE) and support of advanced IP services like differentiated services (DiffServ) [15,16]. Traffic Engineering is a process of controlling traffic flows through a network to optimize resource utilization and network performance. However, it is generally more suitable for macro mobility where scalability is a main issue, whereas in micro mobile MPLS, mobility is the main area of concern [17]. Many MPLS based micro-mobility schemes have been proposed [18, 19, 20, 21]. MPLS faces complexity issues as its domain routers have to run different routing algorithms for giving the best QoS paths. DiffServ and IntServ have also been investigated in the mobility framework. Paper [22] investigates the effect of handoff on quality of mobile nodes in DiffServ network.

In the following sections we will discuss Mobility in IPv6, host and network based protocols along with their classification as micro and macro mobility management protocols. Handover analysis of these protocols will be discussed in the end.

2. MICRO AND MACRO MOBILITY PROTOCOLS

Mobility protocols help in mobile node movement from one network to another while keeping its communication channels alive via alternative routes. When IP Mobility is defined within an access network, it becomes a Local Mobility Management Problem also known as micro mobility. An access network is a collection of fixed and mobile network components belonging to one operational domain and providing access to the internet. The area within which the MN may roam may be restricted, but the overall geographic area might be quite large [8]. The access network gateways act as the aggregation routers. Thus localized

http://journals.uob.edu.bh
mobility implies that there is some administrative management of all the components of the domain defined as local. There is some association between the components as opposed to none in case of a global mobility management scenario. A Global Mobility Management Protocol is a mobility protocol that maintains session continuity when a MN moves from one network to another causing change in the network topology [8]. The global, end-to-end routing of packets is changed during mobility to maintain session continuity. In this case there is no administrative management between the components that are allowed and as such there is no restriction of mobility to be within an access network. A comparison of micro and macro mobility scenario is shown in Fig. 1.

Each access network has a gateway and below it fall all the other routers belonging to the network. If we have two such networks, Network-I (N-I) and Network-II (N-II), a MN moving between these two access networks will fall into the global mobility scenario and such mobility has to be managed by a global mobility protocol like MIPv6, HIP, MOBIKE etc. However if a MN moves between two routers of the same access network it will fall in the domain of local mobility and will be managed by a local mobility protocol like PMIPv6. A router having more than one access point implies that any MN movement between the two access points consists of intra-link mobility. It involves only Layer 2 mechanisms and as such it is also known as Layer 2 mobility. There is no IP subnet configuration necessary once the MN moves between access points of the same router as the link does not change. However some IP signaling may be required [8]. When the MN moves between two access points belonging to two different routers in the same access network, then it becomes a case of micro mobility. Note that in case of R-IIb, there is no intra-link mobility possible. However, it is possible under R-IIa since it has two access points. Any node that moves from access point AP-Ia, AP-Ib, AP-Ic, AP-Id to any access point AP-IIa, AP-IIb, AP-IIc will fall under a global/macro mobility management protocol (scenario V in Fig. 1) and any movement between access points of the same access network will comprise intra-link (scenario I, II, III of Fig. 1) and local/micro mobility management protocol will be required for movement between access points belonging to different access routers as shown in the figure (scenario IV, VI of Fig. 1).

In case of global mobility protocols also known as macro mobility, the MN is reachable even when its globally routable IP address changes. This is done by the home address and care of address mapping pair that is created and is updated at the CN (in case of route optimization) or at the HA or the global mobility anchor point. Since the basic mobility scenario is the same if the MN moves between routers of the same access network or between routers of different access networks, global mobility protocols can substitute for local mobility protocols. However it is not efficient to use global mobility management protocols for local mobility management. Firstly because updating the care of address at the HA, CN or the global mobility anchor point can be time consuming and result in packet loss when packets continue to be sent to the original or the home address of the MN. Secondly update messages involve signaling between the MN and the HA, or MN and the CN, keeping the mobile node occupied for some time. This creates performance overhead for the MN as well as the wireless network. It can impact wireless bandwidth usage and all effect real-time communications.

Another important issue with using global mobility protocol is issue of location privacy [8][23][24]. Privacy in internet is aims at protecting user communication from exposing information about the internet user involuntarily or unintentionally since this information can be used to examine and gather sensitive user data. If the care of address of the mobile node keeps changing, signals need to be exchanged to update the CN, HA or the global mobility anchor point. Traffic analysis will catch these signals and indicate that a particular node in the network is roaming. The change in the IP address will also reveal the location of the MN as IP address reveals the topology of the network and as a result the location of the MN as well. Thus using global mobility management protocols for localized mobility or intra-link mobility has some drawbacks. Therefore need of a localized mobility management protocol arose and gave way to Network - based Localized Mobility Management (NETLMM) [24].

Table I categorizes mobility management protocols as macro or micro mobility protocols.
TABLE I. MOBILE IP PROTOCOLS CLASSIFIED AS MACRO/MICRO AND HOST/NETWORK BASED

<table>
<thead>
<tr>
<th>Protocol</th>
<th>MIPv4</th>
<th>MIPv6</th>
<th>HMIPv6</th>
<th>FMIPv6</th>
<th>PMIPv6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scope</td>
<td>Macro</td>
<td>Macro</td>
<td>Micro/Macro</td>
<td>Macro/Micro</td>
<td>Micro/Macro</td>
</tr>
<tr>
<td>Scope</td>
<td>Host</td>
<td>Host</td>
<td>Host</td>
<td>Host</td>
<td>Network</td>
</tr>
</tbody>
</table>

3. HOST BASED MOBILITY PROTOCOLS

When a MN connects to a network over a wireless interface, it has to acquire the new IP address based on the network topology and send a location registration message to the home agent. Once the MN receives an acknowledgement, its registration is complete. If the MN continues to stay in the same network it keeps on periodically updates the HA about its presence in the network and receives acknowledgements for the same. Whenever a MN moves to a new network, it keeps its new network aware of its presence by sending periodic location update messages. Thus the mobile node is constantly engaged in the process of signaling to the HA in the network in which it enters. This approach in which the MN is engaged in the mobility management is known as Host-based mobility management.

HA has to wait for a Binding Update from the MN before it can send packets to the MN in the new network. A MN might wait for the Binding Acknowledgment before it can send something to the HA for further communication. If Route Optimization is used, then there might be a further delay before communication might actually start between the MN and the CN. These Binding updates need to be authenticated and that might altogether take around 1.5 round-trip times between the mobile node and each correspondent node. Although communication between MN and CN and MN and HA can be carried out in parallel, but further optimizations need to be done to reduce the round trip times. [25] [26] [27]. All these messages exchanged by the MN with other agents in the network and with the CN for address updating consume bandwidth and cause link layer and IP layer delays which might affect the protocols in use. Reducing the delay in the process of handover is essential to the performance improvement of MIPv6 [28].

MIPv6 is a host based macro mobility management and is discussed first. The MN is involved in the signaling process. The delay is long due to this signaling process and as such there is more packet loss in MIPv6. This led to the creation of extended versions of MIPv6, namely HMIPv6 by H.Soliman and FMIPv6 by R.Koodli. A combination of the above two led to the creation of FHMPIPv6 by Hee Young Jung et.al. Tran Cong Hung et.al gave oF-HIPv6 which is an optimized version of FHMPIPv6 [29]. These support micro mobility along with macro mobility hence reduce signaling. Some important extended versions and their performance improvement in comparison with MIPv6 are discussed here.

A. MIPv6: Macro Mobility

In a basic mobility scenario, a MN moves to another network. Once movement is detected, it is followed by a tie up with an agent in the foreign network known as the foreign agent (FA) and acquiring of a new IP address known as the Care-of-Address (CoA). The FA not only helps the MN to inform it’s Home Agent (HA) about its new location, it also serves as an intermediary by accepting packets from HA and forwarding them to MN. Thus home agent helps in maintaining the connection and makes the network switch transparent to the communicating entities. The entity or node communicating with the MN is known as the Correspondent Node (CN). The reply to the CN is send directly from the MN, which leads to triangular routing [30].

When a node changes its network, it has to identify this movement. A delay occurs when a node moves from one network to another. The node figures out this movement by matching its IP address prefix with the prefix of the network. This is known as the movement detection delay ($T_{mdd}$). Once the movement is detected, the node has to wait for a router advertisement. A router advertisement is a message that is sent out periodically by a router to a multicast capable link to announce its availability. If the node receives no such advertisement, the node itself sends a router solicitation to get a router advertisement message immediately. These advertisements and solicitations’ help in router, prefix and parameter discovery as well as address auto configuration [31] [32]. This introduces another kind of delay known as the router advertisement delay. Once the router is recognized and a new IP address obtained in the foreign network, there is an IP address check to ensure that there is no duplication of IP addresses. The time taken is known as the duplicate address detection delay. The router advertisement delay and the duplicate address detection delay together represented as $T_{dad}$ is followed by more delays before communication can be restored. The signals and notifications exchanged between the node and its home network introduces the binding update delay ($T_{bud}$). In certain applications long delays cause packet discarding at the destination [33]. If route optimization is used, then additional time is required to register the new CoA with the CN ($T_{o}$). Thus the whole procedure of movement and address configuration should be aimed to minimize such delays and help in the smooth functioning of the mobility protocol [3]. All these delays comprise handover which is basically a process of terminating the existing connections and setting up new IP connections. The total delay caused is the handover latency. [33]
All messages that report the new position of the mobile node to the home network must be authenticated in order to protect them against remote redirection attacks [4]. The key communication security problems that arise include authentication and key establishment between the nodes that are communicating with each other. Integrity, replay and confidentiality protection of the protocols used for such communication is also an important security concern and is provided by Internet Protocol Security (IPSec) [34].

MIPv6 has advantages over MIPv4 because of the additional features available in IPv6 [32]. MIPv6 also has the additional advantage of having a large pool of IP addresses available to it because of the 128 bit address space of IPv6. IPv6 also known as the Next Generation Network (NGN) has security features that give MIPv6 advantage over MIPv4. Route Optimization is a part of MIPv6 specification; all IPv6 nodes are expected to support it. This was not the case with MIPv4. The other difference is the absence of a foreign agent; the mobile node is a direct point of communication with the home agent. Mobile IPv6 uses two IP addresses per node. One is the home address; the address a mobile node has in its home network. This address is fixed and anyone on the internet can communicate with this node through this address. The other address is the Care-of-Address; the address a mobile node has in the foreign network. It changes as the mobile node moves from one network to another [35]. A home address and a care-of address pair is known as binding. This binding is valid only for a particular interval and needs to be refreshed periodically. It is the responsibility of the mobile node to update the HA with its new CoA [30] [35]. Once this update is received, packets are tunneled to the care of address. This tunneling leads to triangular routing as shown in Fig. 2. No foreign agent is present in this case.

The updates to the HA and the CN are sent through notifications. In IPv6 there are three new procedures known as the Binding Update, Binding Acknowledgement and Binding Request [36]. The CoA is communicated using these notification procedures. The MN can send a Binding Update to a correspondent and later the correspondent can send packets directly to MN, without having HA as an intermediate [36]. This is done using Route Optimization supported in Mobile IPv6. Fig. 3 depicts the notification procedures in Route Optimization. The CN sends packets to the CoA with a routing header. This routing header with the MN’s home address ensures that the exact socket of communication is selected. It also helps in swapping the CoA with the MN’s original address so that at the higher level the connections are maintained [6]. Route Optimization uses Return Routability Procedure [6]. It involves two kinds of checks to ensure that there is a node to which packets can be sent to and accepted from. The Home Address check and the Care of Address check consist of messages that are sent to the Home Agent and the CN respectively.

The communication is carried out using ICMP version 6 (ICMPv6). Router Advertisements, Router Solicitation, Address auto configuration and neighbor advertisements are all carried out using this protocol. Neighbor advertisements are sent out by the home agent in its home network to associate the MN’s IP address with its machine address. This enables the HA to intercept packets destined for MN [36]. Thus MIPv6 supports mobility without having to worry about the presence of agents in other networks. Also, the built in Route Optimization feature removes the dependence on the home network. The extensible headers of IPv6 help in securing all transactions as well as update and acknowledgement messages.
B. HMIPv6

Hierarchical Mobile IPv6 introduces a new node known as Mobility Anchor Point (MAP). MAP can be present anywhere in the hierarchy of routers including access routers belonging to the access networks. Thus when a MN moves, it needs to interact with the MAP instead of the HA which might be lower down in the hierarchy of routers and as such closer to the MN than a HA. This give HMIPv6 an advantage over the MIPv6 protocol by creating a network entity closer to the MN known as the MAP.

HMIPv6 also helps in hiding the location of the MN from the CN and HA. When a MN moves it sends a binding update to the MAP. This is the only BU that it has to send irrespective of the number of nodes it is communicating with. Therefore MAP acts as a local home agent and helps the MN to hide its location from the correspondent nodes and the home agents. The rest of the mobility procedure remains same with Route Optimization being utilized. Security needs to be maintained for all the three scenarios that can arise in HMIPv6. MN to MAP, MN to HA and MN to CN mappings and message exchanges need to be secured. HMIPv6 also allocates a Regional Care of Address (RCoA) to the mobile node. This is done by the MAP. The MN should be aware of this protocol and should be able to receive and process MAP option from the local router. The MN aware of HMIPv6 should also be able to send binding updates with the M flag set. The On-Link Care of Address (LCoA) is configured on the MN’s interface based on the prefix advertised by the default router. This is similar to the CoA of MIPv6, and RCoA is similar to the home address in case of MIPv6. A local binding update is sent by the MN to the MAP to create a mapping between RCoA and LCoA. This is because the RCoA remains the same when a MN moves within the domain of the MAP, but the LCoA changes and hence a binding update is required to be sent to MAP. However if the MN moves out of the domain of MAP, then the MN needs to register the RCoA to the CN and HA as well as create a RCoA and LCoA binding in the new MAP. Thus these two addresses provide both macro and micro mobility support in HMIPv6. The MAP sends packets to the MN without any modification, and as such the MN knows the addresses of the CN, but the CN node is just aware of the RCoA. The packets are forwarded from the RCoA to LCoA by the tunnel [29].

The Operations of HMIPv6 can therefore be classified into four stages. These are the MAP Discovery, MAP Selection, Movement Detection and Binding Updates [24]. MAP Selection is done by the MN through the user of Router Advertisements. This message is created and propagated downwards through all the interfaces by the MAP lying highest in the hierarchy of MAPs. All the MAPs on the way down add their option and spread it downwards. The MAP option comprises of MAP-ID, distance from the highest map, preference, and globalID. GlobalID is used by MN to generate the RCoA. The preference is decreased by one every time a MN chooses to use a particular MAP. The preference of selection is also based on the distance from the MN to the MAP [37]. A MN may choose a distant MAP to avoid re-registration. But the preference field along with the speed of registration with a far MAP will be a criterion for MAP selection as well. A MN can get registered to more than one MAP and as a result have more than one RCoA-LCoA mapping. It can use then use each MAP address for a specific group of correspondent nodes [28]. This utilizes the network bandwidth in an efficient manner.

When a Mobile Node moves between routers that fall under the same MAP, only the LCoA changes and an update is sent to the MAP as shown in Fig. 4. This means that HMIPv6 is handling mobility locally. The HA and the CN communicate with the MN still via the same MAP and are aware of the RCoA of the MN. If the RCoA changes, which implies the MN moves out of the domain of the MAP it is associated with, (as shown in Fig. 4) an update is sent to the HA and the CN regarding the new RCoA as well as a mapping between the LCoA and RCoA is created in the new MAP.

This means that HMIPv6 handles global mobility via the RCoA and local mobility via LCoA. HMIPv6 reduces signaling load outside the MAP domain when handoffs are performed within the local domain and as such handoff performance may be improved and handoff latency reduced. This also reduces packet loss and signaling overhead. However this reduction in signaling is dependent upon the movement of the MN within or outside the MAP domain. There is no
reduction in periodic Binding Updates sent to the MAP [38] [39].

C. FMIPv6

During the process of handover, there is a time period during which the MN is unable to send or receive any packets. As discussed earlier, handover latency is a result of many delays and since there is packet loss in MIPv6 during handover, the extended MIPv6 protocols aim to reduce such delays and improve performance to make the protocols effective for real time communication such as Voice over IP (VoIP). FMIPv6 precisely aims at reducing handover latency [33]. It is based on the idea that the MN is aware of the IPv6 subnet it is going to move to before the actually movement takes place. The access router in the foreign network can buffer all the packets destined for the MN that arrive till it actually gets connected after handover [40]. The ability to immediately send packets from a new subnet link depends on the delays caused due to MN movement as discussed earlier. The "IP connectivity" latency depends upon the delay caused till movement is detected and the configuration of the new CoA, which depends upon $T_{dadd}$, $T_{had}$ and $T_{ro}$. Therefore receiving packets at the new address is dependent upon the Binding Update latency as well as the IP connectivity latency [31] [33].

A MN in its home network has address PCoA (Previous Care of Address) and is connected to the access router known as the Previous Access Router (PAR). When it moves to the new network, it connects with the New Access Router (NAR) and acquires the New Care of Address (NCoA). Fast handover consists of three steps: Handover initiation, tunnel establishment and packet forwarding [41]. FMIPv6 uses Router Solicitation for Proxy Advertisement (RtSolPr) and Proxy Router Advertisement (PrRtAdv) for fast handover. A MN is in its home network it can detect the presence of other access points and can ask its access router for the subnet information of all the access routers that it can detect. Handover is initiated when a MN sends an RtSolPr message to the PAR to indicate that it wants to perform a fast handover to a new AR. This message consists of the link layer address of the new point of attachment that is discovered from the NAR’s beacon message. Thus RtSolPr is a message sent from the MN to the PAR requesting information about the other access points under whose influence it might be in while still being in the home network. The PAR replies with a PrRtAdv that provides the MN information about the neighbouring links and both of these messages together help in expedited movement detection. A tuple (AP-ID, AR-Info) contains an access router’s (AR) L2 and IP addresses, and the prefix valid on the interface to which the Access Point (identified by AP-ID) is attached. The triplet (Router’s L2 address, Router’s IP address, Prefix) is the AR-Info field. This is the tuple that the MN receives when it moves to a new access point with AP-ID. MN finds out the rest of the information from the AR-Info field of the tuple, thus helping in expedited movement detection. MN also forms an NCoA while it is still connected to PAR. Thus this address can be used immediately once movement is detected and address configuration delay is as such reduced helping in making the overall handover process faster. MN sends a Fast Binding Update (FBU) to the PAR using this NCoA and receives a Fast Binding Acknowledgement (FB-ACK) to indicate success. If it is feasible for MN to send the FBU from the PAR’s link, then that should be preferred. Otherwise it should be sent immediately after the NAR has been detected [33] [41].

A tunnel between the PCoA and the NCoA is created when a PAR sends a Handover Initiation (HI) message to NAR and it replies with a Handover Acknowledgement (Hack). After the tunnelling phase is over, packet forwarding starts. PAR begins tunnelling packets arriving for PCoA to NCoA. The tunnel remains active until the MN completes the Binding Update with its correspondents. Forwarding support for PCoA is provided through a reverse tunnel between the MN and the PAR since correspondent nodes have to be updated with a Binding Cache entry that has the NCoA. MN sends a Fast Neighbour Advertisement (F-NA), to start the packet flow from NAR to itself [33] [41]. Fig. 5 shows the mechanism for fast handover in MIPv6.

![Figure 5. Scenario for fast handover in MIPv6](http://journals.uob.edu.bh)
Handover can be improvised or optimized using two modes in which FMIPv6 can operate in. These classify as Predictive Handover and Reactive Handover. FMIPv6 predictive mode allows it to fully benefit from all FMIPv6 optimizations and has been described above. Reactive handover mode used when a node suddenly loses its connection with its current router or access point. It means that the MN cannot foresee a handover and as such is able to react only when it is already in the process of handover. Thus it is known as the reactive mode. Once the MN is in NAR’s link, a FBU is sent and is usually encapsulated in the FNA. The NAR forwards the FBU to the PAR and the PAR starts the tunnelling phase after receiving the FBU. [33][40]. FMIPv6 also allows the AR to send an unsolicited PrRtAdv to the MN including the tuple for any neighbouring access networks. When a MN receives such a message it starts predictive handover to the network mentioned in the tuple. This is a network initiated handover and may be used for purposes of load sharing [33] [40]. The signals exchanged in both the predictive and reactive mode is shown in Fig.6. Since FMIPv6 takes care of many things while still in the home network, its handover is faster than a MIPv6 managed mobility network in which a MN moves to a foreign network and then acquires a new address and updates its HA to start triangular routing or route optimization. This protocol takes care of many things while still in the home network, so if it moves within a domain or outside a domain, both the cases of micro and macro mobility will be taken care of.

4. NETWORK BASED MOBILITY PROTOCOLS

Network based mobility protocols aim to keep the MN unaware of the process of mobility while retaining its connections during its movement from one network to another. Internet Engineering Task Force (IETF) is working on many global mobility protocols and as such in future there will be many Mobility Management Protocols available in a wireless network. This will give the MN an option to select one of the available mobility protocols. Depending on what mobility protocol the MN chooses, it will have to make some software stack changes, like deploy and implement protocol specifications so that it can retain its connections when it is mobile. If the MN is part of the mobility management, there is the burden of deployment and implementation on the host (MN) and increase complexity on the MN. To reduce the onus from the MN to install the stack-software compliant to a particular type of mobility management, new protocols are being introduced to shift the burden of mobility management to the network only, sparing the host from the trouble of changing its stack software every time it wants to switch between the different available protocols.

Network based mobility protocols have been more in focus as they can give host the liberty to select a protocol of its choice, but it will also bring more MNs within the mobility circle [8][42]. In Network-based mobility management the network handles the mobility management on behalf of the MN; thus the MN is not required to participate in any mobility-related signalling. Network Based Localized Mobility Management (NETLMM) emerged because of the problems that existed with Host based mobility management protocols. It provides as interoperable, uniform localized mobility management protocol that is extendible to topologically large networks, without requiring host stack contribution for localized mobility management. It aims at providing fault tolerance, robustness, interoperability, scalability, and minimal specialized network equipment. Other goals of NETLMM include handover performance Improvement and reduction in handover related signalling volume. It solves the problem of location privacy as exists in global mobility protocol, as it is a localized mobility management protocol. It supports unmodified mobile nodes and support for IPv4 and IPv6. Localized mobility management also aims at reusing existing protocols when sensible. Also the choice of localized mobility management should not curb or be constrained by the choice of the global mobility protocol. [8][43].

NETLMM defines an access network with a Mobile Access Gateway (MAG) and a Localized Mobility
Anchor (LMA). LMA is a router that maintains a collection of host routes and associated forwarding information for mobile nodes falling under its domain. It is just like the HA in MIPv6 with extended functionalities [44]. It manages the IP node mobility together with the MAG. When a MN moves around in the localized mobility management domain, its data routing is anchored at the LMA. MAG is the network entity that is responsible for the mobility management on behalf of the MN. All the signalling related to the MN is communicated by the MAG to the LMA. In fact MAG is the connection of the MN to the localized mobility management domain [44].

A. PMIPv6

PMIPv6 is a network based mobility management protocol standard and has been ratified by the NETLMM working group of IETF. The MN involvement can be removed from mobility scenario completely by involving a network node (MAG) and a home agent (LMA) in the signalling process. Security associations are set up between LMA and MAG and authorization for sending signals like the binding updates on behalf of the mobile node to ensure there are no security issues. A network in which mobility management of a MN is managed using PMIPv6 is known as the Proxy Mobile IPv6 Domain and it consists of LMA’s and MAG’s. An LMA in a PMIPv6 domain is a home agent with additional functionalities that serves as the anchor point for the MN. It manages the mobile nodes binding state and topologically it is the anchor point for the mobile nodes possible home addresses. MAG is an access router that tracks the mobile nodes movement to and from the access link and informs the LMA regarding the movement. LMA has an address configured on its interface known as the LMA Address (LMAA). It is the end point of the bi-directional tunnel between it and the MAG. Proxy Care –of –Address (PCoA) forms the other end of this bi-directional tunnel. This address is configured on the outer interface of the MAG and this serves as the care of address of the MN and is used in the Binding Cache entry for the MN [45].

A prefix advertised by the MAG in the PMIPv6 domain for a MN is known as the Mobile Node Home Address Prefix (MN-HNP). The MN always configures its address from this prefix that is anchored at the LMA. The address used by the MN for communication is the Mobile Node Home Address (MN-HoA). The LMA is aware of the MN-HNP and not the exact MN-HoA that is configured on the MN. The MN can use this address at all attachment points in the PMIPv6 domain. [44]. A MN can connect to the same PMIPv6 domain through multiple interfaces and can use these interfaces simultaneously. Such a MN is known as a multi-homed MN. MN Identifier (MN-NAI) [31] is an identifier that is used to perform the authentication procedures with LMA and MAG. MN Link Layer Identifier (MN-LL-Identifier) identifies the attached interface of a MN. Sometimes it is generated by the MN and conveyed to the MAG. The network manages a database known as the policy profile that contains information about different parameters of the MN that the MAG and LMA may require to provide mobility related services to the MN [45].

Like in MIPv6, PMIPv6 involves a Proxy Binding Update message (PBU) that the MAG sends to the LMA on behalf of the MN. A Proxy Binding Acknowledgement (PBA) is sent from the LMA to MAG in response to PBU. The following Fig. 7 shows the identities in the PMIPv6 network. Fig. 8 depicts the message flow. When a MN enters a PMIPv6 domain for the first time and attaches itself to an access link of a MAG in the PMIPv6 domain. The MAG registers the location of the MN to the MN’s LMA on behalf of the MN after determining if it is authenticated to enter the PMIPv6 domain. This is done by using the PBU message. Once LMA receives the PBU, it allocates a H-NP and creates a binding cache entry for the MN and replies with the PBA message. This establishes the bi-directional tunnel for packet delivery. The tunnel can be shared by all the MN’s connected to the same MAG and LMA. When a MN moves away to another MAG, the tunnel is released only after the lifetime of the tunnel expires (if it is shared) or if there are no more MN’s sharing it [46].

When a MN moves from one MAG to another, the MN-HNP does not change for the MN. The MN considers PMIPv6 domain as the domain that has the same home link. The MN sends MN-Identifier messages to the new MAG for authentication. The MAG obtains the mobile nodes profile after the layer 2 handover is completed so that it can notify its movement by policy store that consists of an AAA server where the three A’s stand for Authentication, Authorization and Accounting.
The profile consists of information about MN-ID, LMA accepted address modes, and roaming policies for providing network based mobile services. After obtaining the profile of the MN from the policy store, the MAG sends router advertisement messages with a home network prefix to the MN if the profile contains a MN’s home network prefix. This MAG sends a PBU to the LMA for registering the MN’s current location information. LMA checks in its binding cache if it has an entry for the MN-ID. If it is not present an entry is created and a tunnel is set up between the LMA and the new MAG and a PBA is sent to the new MAG with MN’s home network prefix options [42]. Otherwise a tunnel is set up between the new MAG and the LMA and the LMA starts forwarding packets meant for the MN through the tunnel.

Thus in case of PMIPv6, there is no change of the MN-HNP and the domain emulates the home network like the previous MAG. Thus the MN can be mobile without any mobility stack as the mobility management is restricted to the network itself. This is the advantage of PMIPv6 over host based protocols like FMIPv6 or HMIPv6. Also, within a MAG it is the case of local mobility and outside it when the MN has to re-register with the LMA, it becomes the case of global mobility, Hence, PMIPv6 takes care of both micro and macro mobility. Studies have shown that the handover latency of PMIPv6 compared to other mobility protocols is considerably lower [47].

The handover performance of protocols is discussed in the next section.

5. Handover Performance Comparison

To compare handover performance of the protocols discussed above, let us have a look at the handover procedure in each of the above mobility management protocols. Handover is a process of terminating the existing connections and setting up new IP connections. The time taken during this process is the handover latency [33]. During the process of handover, there is a time period during which the MN is unable to send or receive any packets. The main aim is to improve performance with respect to the handover latency and minimize packet loss. A similar analysis is done in [47]. In [47], handover latency is defined as the time that elapses between the moment when L2 handover completes at the AP and the moment MN receives the first packet after moving to the new point of attachment.

In the MIPv6 protocol, when a MN moves from one network to another, before communication is re-established signals are exchanged. Fig. 9 shows the signals exchanged in MIPv6.

The total time \( T_M \) required for the handover is the sum of the delays as described earlier.

\[
T_M = T_{mdd} + T_{dadd} + T_{bud} \quad (i)
\]

If Route Optimization is used then the total handover time is equal to:

\[
T_{MRO} = T_{mdd} + T_{dadd} + T_{bud} + T_{ro} \quad (ii)
\]

In all cases of mobility, the MN moves from the Home Network, away from its home agent to the foreign network. The foreign network can have different components depending on what protocol we are discussing. To perform an analysis of handover in other Mobility protocols, we consider a general network model as shown in Fig.10.
Figure 10. Network Model for Analysis of Handover

The figure shows three cases for MIPv6/FMIPv6, HMIPv6 and PMIPv6 respectively. In case of MIPv6 and FMIPv6, it is simply the foreign network with no foreign agent as FA is removed in case of MIPv6. For HMIPv6, the communication will go from MN, to the Access Point, from Access Point to the Access Router and then from Access Router to the Mobility Anchor Point (MAP). Similarly, in case of PMIPv6, the MN is assumed to move within or outside the MAG domain. Depending on the type of movement, MAG and LMA exchange messages or LMA registers a new MN after receiving a request from the corresponding MAG. The HA network and the network to which the MN moves can fall within the same cloud, but they will differ in their places in hierarchy. For explanation purposes, we separate out the clouds. The arrows depict the various delays when messages are exchanged. These are explained in Table II.

Table II. Delays and Their Descriptions

<table>
<thead>
<tr>
<th>Delay</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_{\text{max}}$</td>
<td>Delay between MN and the Access Point (AP), when packet is sent through a wireless link</td>
</tr>
<tr>
<td>$T_{\text{tar}}$</td>
<td>Delay between AP and AR/MAG/Router (AP and AR in case of HMIPv6, AP and MAG in case of PMIPv6, AP and the Router in case of MIPv6 and FMIPv6)</td>
</tr>
<tr>
<td>$T_{\text{ary}}$</td>
<td>Delay between AR/MAG and MAP/LMA. (AR and MAP in case of HMIPv6, MAG and LMA in case of PMIPv6)</td>
</tr>
<tr>
<td>$T_{\text{arh}}$</td>
<td>Delay between AR/MAG/Router and HA for HMIPv6, PMIPv6 and MIPv6/ FMIPv6 respectively.</td>
</tr>
<tr>
<td>$T_{\text{arc}}$</td>
<td>Delay between AR/MAG/Router and CN not via HA for HMIPv6, PMIPv6 and MIPv6/ FMIPv6 respectively.</td>
</tr>
<tr>
<td>$T_{\text{hc}}$</td>
<td>Delay between HA and CN.</td>
</tr>
</tbody>
</table>

A binding update from MN to HA will therefore incur a delay of:

$$T_{\text{max}} + T_{\text{tar}} + T_{\text{arh}}$$  (iii)

A binding update from MN to CN will therefore incur a delay of:

$$T_{\text{max}} + T_{\text{tar}} + T_{\text{arc}}$$  (iv)

Considering symmetry of signals the above two equations can be doubled to calculate the total time taken for Binding Update and Acknowledgement between MN to HA and MN to CN respectively. To use the return routability procedure, from the figure the total time required one way is equal to:

$$T_{\text{max}} + T_{\text{tar}} + T_{\text{arh}} + T_{\text{hc}}$$  (v)

Therefore the total time delay for MIPv6 can be calculated by substituting (iii), (iv) and (v) in (ii). We get:

$$T_{\text{MRO}} = T_{\text{mdd}} + T_{\text{dadd}} + 4(T_{\text{max}} + T_{\text{tar}}) + 2(T_{\text{arh}} + T_{\text{arc}}) + 2(T_{\text{max}} + T_{\text{tar}} + T_{\text{arh}} + T_{\text{hc}})$$

Which equals:

$$T_{\text{MRO}} = T_{\text{mdd}} + T_{\text{dadd}} + 6(T_{\text{max}} + T_{\text{tar}}) + 2T_{\text{arh}} + 2T_{\text{arc}} + 2T_{\text{hc}}$$

In HMIPv6, the introduction of MAP reduces the BU messages exchanged between the HA and the MN. Therefore the handover delay of HMIPv6 consists of the factors $T_{\text{mdd}}$, $T_{\text{dadd}}$ and $T_{\text{bud}}$, where $T_{\text{bud}}$ is the new binding update delay that is smaller than its counterpart in MIPv6. Since MAP is closer to the MN than a HA, HMIPv6 gets the advantage over MIPv6 protocol by having a network entity closer to the MN and also hiding the location of the MN from the CN and HA. Therefore, from (i)

$$T_{\text{H}} = T_{\text{mdd}} + T_{\text{dadd}} + T_{\text{bud}}$$  (vi)

From the figure, $T_{\text{bud}}$ equals $2(T_{\text{max}} + T_{\text{tar}} + T_{\text{ary}})$. After substituting the $T_{\text{bud}}$ value in (vi), the final delay is equal to:

$$T_{\text{H}} = T_{\text{mdd}} + T_{\text{dadd}} + 2(T_{\text{max}} + T_{\text{tar}} + T_{\text{ary}})$$

Fast handover consists of three steps: Handover initiation, tunnel establishment and packet forwarding. MN also forms an NCoA while it is still connected to...
PAR. Thus this address can be used immediately once movement is detected and address configuration delay is as such reduced helping in making the overall handover process faster. MN sends a Fast Binding Update (FBU) to the PAR using this NCoA and receives a Fast Binding Acknowledgement (FB-ACK) to indicate success. PAR begins tunneling packets arriving for PCoA to NCoA. The tunnel remains active until the MN completes the Binding Update with its correspondents. Thus when MN sends a Fast Neighbor Advertisement (F-NA), to start the packet flow from NAR to itself it cause 2(T_{max} + T_{sar}) delay. Fig. 5 shows this mechanism for fast handover in MIPv6.

Thus, in FMIPv6, the CoA configuration and the duplicate address detection is done before it disconnects from the link to the network it was previously in. Therefore T_{std}' and T_{dadd}' is removed from T_M in (i) as it is the time taken to configure the CoA and check for duplicity which is done beforehand in FMPIPv6. Therefore the total handover latency time (T_F) of FMIPv6 can be expressed as :

\[ T_F = T_{bud} \text{ where } T_{bud} = 2(T_{max} + T_{sar}) \]

Therefore,

\[ T_F = 2(T_{max} + T_{sar}) \]

Fast handover mechanism has an advantage over MIPv6 and HMIPv6 because the term T_{std}' and T_{dadd}' is missing from the expression for total handover latency. If we use F-HMIPv6 protocol, then we get rid of the term T_{std}' and T_{dadd}' from the total time as well as get the reduced T_{bud}' time for binding updates making the total time for a F-HMIPv6 protocol as follows:

\[ T_F = T'_{bud} = 2(T_{max} + T_{sar} + T_{ary}) \]

So the best handover latency is achieved when FMIPv6 and HMIPv6 are used together as is done in F-HMIPv6 protocol. The result is signalling load reduction, improvement in latency delay and less packet losses apart from helping the handover process by pre-configuration of CoA.

For Network based mobility management protocol like PMIPv6, the handover latency is calculated as follows. After obtaining the profile of the MN from the policy store, the MAG sends router advertisement messages with a home network prefix to the MN if the profile contains a MN’s home network prefix. This takes T_{max} + T_{sar} delay. This MAG sends a PBU to the LMA for registering the MN’s current location information. LMA checks in its binding cache if it has an entry for the MN-ID. If it is not present an entry is created and a tunnel is set up between the LMA and the new MAG and a PBA is sent to the new MAG with MN’s home network prefix options. The handover latency from these signal exchanges is equal to the time to send PBU from MAG to LMA and receive the PBA from LMA to MAG plus the time required for the packet forwarded by the LMA to reach the MAG. According to the figure it equals 2(T_{ary} + T_{max} + T_{sar}).

\[ T_F = 2(T_{ary} + T_{max} + T_{sar}) \]

Table. III summarizes the handover latency in macro and micro mobility protocols. We see micro mobility protocols have a reduced delay compared to macro mobility protocol.

**6. Conclusion and Further Work**

Micro and macro mobility has been overviewed with respect to the protocols of the next generation internet. MIPv6 leading all of them has been improvised by its extensions namely HMIPv6, FMIPv6 etc, all of whom are host based. NETLMM has come up to set up a mobility management network without involving the host or the MN in mobility related signalling. This increases the scope of nodes that can participate in mobility and call themselves mobile with respect to internet connectivity. Our last section outlines the signalling involved in handover and the total handover latency in each of the protocols discussed. These protocols can be used in conjunction with latest technologies and improved for QoS by using these protocols in a MPLS framework. Our future work involves studying issues with these protocols and the problems faced when using these different protocols within MPLS domain to improve their QoS. We are also in the process of using Network Simulator to study these protocols so that we can emulate larger networks. Test bed development is also in progress.

**TABLE III. HANDOVER COMPARISON IN MICRO AND MACRO MOBILITY PROTOCOLS**

<table>
<thead>
<tr>
<th>Protocol</th>
<th>Handover Analysis</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIPv6</td>
<td>T_{MN0} = T_{sat} + T_{dadd} + T_{sat} + T_{w}</td>
<td>Lots of signaling, binding updates leading to more delay.</td>
</tr>
<tr>
<td></td>
<td>where T_{sat} = (T_{max} + T_{sar}) + 2 (T_{std} + T_{w})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and T_{w} = 2(T_{sat} + T_{sar} + T_{dadd} + T_{std})</td>
<td></td>
</tr>
<tr>
<td>HMIPv6</td>
<td>T_{MN} = T_{sat} + T_{sat} + T_{sat}</td>
<td>MAP reduces sending BU’s to the HA which can be far away. Thus a</td>
</tr>
<tr>
<td></td>
<td>where T_{sat} = 2(T_{max} + T_{sar} + T_{ary})</td>
<td>reduced T_{sat}</td>
</tr>
<tr>
<td>FMIPv6</td>
<td>T_{FN} = T_{sat} + T_{sat} + T_{sat}</td>
<td>No movement and duplicate address detection. Only Binding Updates.</td>
</tr>
<tr>
<td></td>
<td>where T_{sat} = 2(T_{max} + T_{sar} + T_{ary})</td>
<td></td>
</tr>
<tr>
<td>Fast HMIPv6</td>
<td>T_{HF} = 2(T_{max} + T_{sar})</td>
<td>No movement and duplicate address detection. Reduced Binding Updates</td>
</tr>
<tr>
<td></td>
<td>where T_{sat} = 2(T_{max} + T_{sar} + T_{ary})</td>
<td>due to introduction of MAP.</td>
</tr>
<tr>
<td>PMIPv6</td>
<td>T_{FN} = 2(T_{ary} + T_{max} + T_{sar})</td>
<td>Network based protocol, reduced signaling as MN’s likely to move within</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LMA domain.</td>
</tr>
</tbody>
</table>
REFERENCES


[27] Li Yun, ZHAO Yi-sheng, LIU Qi-Lie, WEN Feng, “Performance Research of MIPv6 and Extended Protocols in the Process of Handover,” IEEE Xplore.


[29] Xavier Perez Costa and Marc Torrent Moreno, “A Performance Study of Hierarchical Mobile IPv6 from a System Perspective”.


Shaima Qureshi has received her B.E (Hons.) Computer Science degree from BITS Pilani, India in 2004. She completed her M.S in Computer Science from Syracuse University, NY, USA in 2006. She is currently pursuing her Ph.D from NIT Srinagar and working as an Assistant Professor in the same Institute since 2008. She has been guiding B.E student projects and has a number of publications to her credit. Prior to joining the academic field, she worked as a Senior QA Engineer for two years in the software industry in USA. Her areas of research include Algorithms, Operating Systems and Computer Networks.

A. H. Mir has done his B.E in Electrical Engineering with specialization in Electronics and Communication Engineering (ECE) from REC Srinagar (J & K) India in 1982. He did his M.Tech in Computer Technology and PhD both from IIT Delhi in the year 1989 and 1996 respectively. He was Chief Investigator of Ministry of Communication and Information Technology, Govt. of India project: Information Security Education and Awareness (ISEA). He has been guiding PhD and M.Tech thesis related to the area of Security and other related areas. He has a number of International publications to his credit. Presently he is working as Professor in the Department of Electronics and Communication Engineering at NIT Srinagar, India. His areas of interest are Biometrics, Image Processing, Security, Wireless Communication and Networks.