5 Performance Evaluation

This chapter evaluates the performance of the GMP compared to the MIP, HMIP and FMIP individual performances. We study the packet loss and the latency to restore the downstream and upstream of packets during intra-domain and inter-domain handover time.

We consider an hipotetical network environment with five different domains interconnected by a backbone. One of the domains is restricted to the stationary CN and the others are visited by the MN during a session connection time. Each domain visited by the MN is composed of a border router (MAP entity) and three ARs connected to it. Each AR is connected to an IEEE 802.11b access point (AP) that provides wireless network access to the MN. All APs are placed in geographical positions that provide the necessary intersection between cells to make seamless handover possible. Figure 18 illustrates the network environment scenario used for the evaluation.

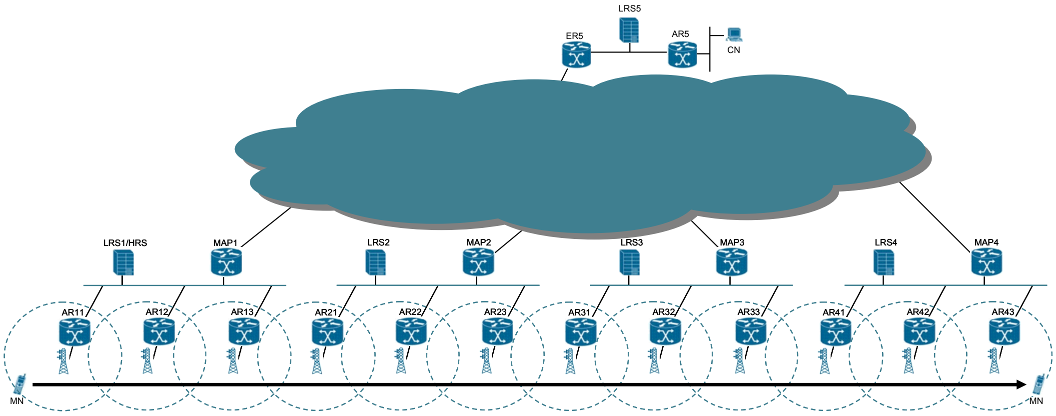

Figure 18 – Network environment scenario used for the evaluation

We also assume that the Media Independent Event Service (MIES) is available and ready to notify the upper layers about the L2-GoingDown, L2-

Down and L2-UP triggers. So, the MN doesn't need to wait for the Router Advertisement message of the new AR to detect an L3 handover. The MN will initiate the L3 handover when the L2-GoingDown trigger is detected. Once notified about the L2-UP trigger, the MN will use the stateless auto-configuration mechanism to avoid contacting any entity to obtain a new co-located CoA. These measures contribute to reduce the handover process duration.

In order to analyze the performance of the mobility management of the protocols, we need to model the transmission delay of the signaling messages. Depending on the protocol, the signaling messages go through wireless links and wired links or just through wired links. Also, some messages are local to a domain and others are originated at a domain and forwarded by intermediary hops to another domain. Table 5 introduces the parameters that will be used in the analysis.

Table 5 – Parameters used in the analysis

Parameters	
T_s	Average session connection time
T_r	Average cell resident time
N_m	Average number of movements during a session (i.e., $\mathrm{N_m} = T_s/T_r$)
S_u	Average size of a signaling message
$D_{x\text{-}y}$	Average number of hops between x and y
B_w	Bandwidth of the wired link
L_w	Latency of the wired link (propagation delay and link layer delay)
B_{wl}	Bandwidth of the wireless link
L_{wl}	Latency of the wireless link (propagation delay and link layer delay)
P_t	Routing table lookup and processing delay
R_d	Downstream transmission rate (packet transmission rate)
R_u	Upstream transmission rate (packet transmission rate)
TL2	Time interval for L2-handover (starts at L2-Down and ends at L2-UP)

Let $T_{wlw}(S,D_{x\text{-}y})$ denote the transmission delay of a message of size S sent from x (always an MN) to y via the wireless and wired links, as suggested by [16]. $T_{wlw}(S,D_{x\text{-}y})$ can be expressed as follows:

$$T_{wlw}(S,D_{x-y}) = \left(\frac{S}{B_{wl}} + L_{wl}\right) + D_{x-y} \times \left(\frac{S}{B_{w}} + L_{w}\right) + \left(D_{x-y} + 1\right) \times P_t$$

Let $T_{ww}(S,D_{x\text{-}y})$ denote the transmission delay of a message of size S sent from x (never an MN) to y via the wired links. $T_{ww}(S,D_{x\text{-}y})$ can be expressed as follows:

$$T_{ww}(S,D_{x-y}) = (D_{x-y} + 1) \times \left[\left(\frac{S}{B_{w}} + L_{w}\right) + P_t\right]$$

5.1 Packet loss during a session

Packet loss (*Pkt_Loss*) during a session is defined as the sum of packets lost during all handovers procedures while the MN is receiving or sending data packets. To reduce packet loss, some buffering mechanisms may be used by both the MN and the AR to store in-flight packets, but these mechanisms usually require more signaling messages to be added to the protocols. Then, not all of the protocols support buffering control of packets during handover time. This is the case of both the MIP and the HMIP. The FMIP and the GMP use signaling messages to enable the buffering and forwarding of in-flight packets from the PAR to the NAR. We will denote the buffer size requirement for downstream and upstream flows as a function of the packet transmission rate in section 5.4 . For now, we assume that the size of the buffer is appropriate to store the in-flight packets. We also assume that the MN is able to store all of its upstream packets and no packets are lost in this direction.

So, to denote the total packet loss for the MIP and HMIP, we consider that in-flight packets are lost during the whole L2-handover and L3-handover. The L2-handover starts when the L2-Down trigger is detected and ends when the L2-UP trigger is detected. The L3-handover starts right after the L2-handover is completed and is composed of procedures that depend on the protocol. For the MIP, these procedures are: (i) discovering of the network prefix (Router Solicitation/Router Advertisement); (ii) registration at the HA; and (iii) establishment of the bi-directional tunnel between the MN and the HA. For the HMIP, the procedures depend on the type of L3-handover. For the intra-domain handover, the procedures are: (i) discovering of the network prefix (Router Solicitation/Router Advertisement); (ii) registration at the MAP; and (iii) establishment of the bi-directional tunnel between the MN and the MAP. For the inter-domain handover, besides the procedures of the intra-domain handover, more two procedures are executed by the MN: (i) registration at the HA; and (ii) establishment of the bi-directional tunnel between the HA and the MAP. The total packet loss for the MIP and HMIP can be expressed as follows:

$$\begin{aligned} Pkt\_Loss(MIP) = [&TL2 + 2 \times T_{wlw}(S_u, D_{MN-AR}) + \\ &+ (2+3+3) \times T_{wlw}(S_u, D_{MN-HA})] \times R_d \times N_m \end{aligned}$$

$$\begin{aligned} Pkt\_Loss(HMIP_{Intra}) = [&TL2 + 2 \times T_{wlw}(S_u, D_{MN-AR}) + \\ &+ (2+3+3) \times T_{wlw}(S_u, D_{MN-MAP})] \times R_d \times N_m \end{aligned}$$

$$\begin{aligned} Pkt\_Loss(HMIP_{Inter}) = [&TL2 + 2 \times T_{wlw}(S_u, D_{MN-AR}) + \\ &+ (2+3+3) \times T_{wlw}(S_u, D_{MN-MAP}) + \\ &+ 2 \times T_{wlw}(S_u, D_{MN-HA}) + \\ &+ (3+3) \times T_{ww}(S_u, D_{HA-MAP})] \times R_d \times N_m \end{aligned}$$

In the case of the FMIP and the GMP, in-flight packets are lost till the buffering and forwarding mechanisms are initiated. The initialization of these mechanisms is executed as part of the L3-handover, based on the mode of operation of the protocols. Usually, the L2-GoingDown trigger is used to initiate the L3-handover before the L2-handover. When the anticipated mode is used, the buffering and forwarding mechanisms are initiated before the L2-handover. Otherwise, in the reactive mode, these mechanisms are initiated after the L2-handover.

For the anticipated mode of the FMIP, these mechanisms are initiated when the FBU message, sent by the MN, is received by the PAR before the L2-handover. So, packets will be lost only if the MN is not able to send the FBU message before the L2-handover and, then, it will need to switch to operate in reactive mode. In this circumstance, after the L2-handover, the MN sends the FNA message (encapsulating the FBU) to the NAR over the wireless link and the NAR sends the FBU message to the PAR over the wired links to initiate the forwarding mechanism. So, the total packet loss for the FMIP can be expressed as follows:

$$\begin{aligned} &Pkt\_Loss(FMIP_{ant}) = 0 \\ &Pkt\_Loss(FMIP_{react}) = [TL2 + T_{wlw}(S_u, D_{MN-NAR}) + \\ &\qquad + 2 \times T_{ww}(S_u, D_{NAR-PAR})] \times R_d \times N_m \end{aligned}$$

The initialization of the buffering and forwarding mechanisms of the GMP is very similar to the procedure executed by the FMIP. For the anticipated mode of the GMP, the buffering mechanism is initiated when the *C2N_HandoverInitiate* message, sent by the MN, is received by the PAR before the L2-handover starts (L2-Down trigger is fired). So, the packets are lost only if the MN is not able to send the *C2N_HandoverInitiate* message before the L2-handover starts. Under this circumstance, like FMIP does, the GMP switches to operate in reactive mode. This mode starts after the L2 handover is complete, when the MN becomes able to send the *C2N_HandoverFinish* message to the NAR over the wireless link. In turn, the NAR sends the *N2N_HandoverIndication* message to the new LRS (NLRS) which relays the message to the previous LRS (PLRS). The PLRS relays the same message to the PAR which finally sends the *N2N_Handover Acknowledge* message to the NAR. To complete this phase, the NAR also exchanges a pair of messages with the PAR over the wired links to request the packet forwarding. So, the total packet loss for the GMP can be expressed as follows:

$$\begin{aligned} &Pkt\_Loss(GMP_{ant}) = 0 \\ &Pkt\_Loss(GMP_{react}) = [TL2 + T_{wlw}(S_u, D_{MN-NAR}) + T_{ww}(S_u, D_{NAR-NLRS}) + \\ &\qquad + T_{ww}(S_u, D_{NLRS-PLRS}) + T_{ww}(S_u, D_{PLRS-PAR}) + \\ &\qquad + 3 \times T_{ww}(S_u, D_{NAR-PAR})] \times R_d \times N_m \end{aligned}$$

5.2 Average handover latency to restore the downstream flow during a session

Average handover latency to restore the downstream flow (*Avg_Lat_Down*) during a session is defined as the average time of all handovers of a session to restore the MN's ability to receive packets of the ongoing downstream. To reduce handover latency, some efforts were made to enhance the L2-handover and the L3-handover. Here, we focus on L3-handover analysis and assume a constant value TL2 to express the L2-handover latency.

We assume that the routing optimization is used to reduce the latency of the packets sent/received by the MN to/from the CN. So, in order to restore the downstream, the MN should execute the return routability procedure before sending the BU message to the CN. This is the case of the MIP and $HMIP_{inter}$. So, the latency to restore the downstream flow for the MIP and the $HMIP_{inter}$ can be expressed as follows:

$$
\begin{aligned}
Lat\_Down(MIP) = {} & TL2 + 2 \times T_{wlw}(S_u, D_{MN-AR}) + \\
& + (2+3+3+2) \times T_{wlw}(S_u, D_{MN-HA}) + \\
& + 2 \times T_{ww}(S_u, D_{HA-CN}) \\
& + (2+2) \times T_{wlw}(S_u, D_{MN,CN})
\end{aligned}
$$

$$
\begin{aligned}
Lat\_Down(HMIP_{inter}) = {} & TL2 + 2 \times T_{wlw}(S_u, D_{MN-AR}) + \\
& + (2+3+3) \times T_{wlw}(S_u, D_{MN-MAP}) + \\
& + (2+2) \times T_{wlw}(S_u, D_{MN-HA}) + \\
& + (3+3) \times T_{ww}(S_u, D_{HA-MAP}) + \\
& + 2 \times T_{ww}(S_u, D_{HA-CN}) + \\
& + (2+2) \times T_{wlw}(S_u, D_{MN-CN})
\end{aligned}
$$

Because of the hierarchical characteristic of the $HMIP_{intra}$, the MN does not need to execute the return routability procedure to restore the downstream. This stream is restored just after the bi-directional tunnel is established between the MN and the MAP. So, the latency in this case can be expressed as follows:

$$
\begin{aligned}
Lat\_Down(HMIP_{intra}) = {} & TL2 + 2 \times T_{wlw}(S_u, D_{MN-AR}) + \\
& + (2+3+3) \times T_{wlw}(S_u, D_{MN-MAP})
\end{aligned}
$$

Because of the buffering and forwarding mechanisms of both the FMIP and GMP, for these protocols, the MN also does not need to execute the return routability procedure to restore the downstream. The latency in these cases can be expressed as follows:

$$
Lat\_Down(FMIP_{ant}) = TL2 + T_{wlw}(S_u, D_{MN-AR})
$$

$$
\begin{aligned}
Lat\_Down(FMIP_{react}) = {} & TL2 + T_{wlw}(S_u, D_{MN-NAR}) + \\
& + 2 \times T_{ww}(S_u, D_{NAR-PAR})
\end{aligned}
$$

$$
Lat\_Down(GMP_{ant}) = TL2 + T_{wlw}(S_u, D_{MN-AR})
$$

$$
\begin{aligned}
Lat\_Down(GMP_{react}) = TL2 + T_{wlw}(S_u, D_{MN-NAR}) + T_{ww}(S_u, D_{NAR-NLRS}) + \\
+ T_{ww}(S_u, D_{NLRS-PLRS}) + T_{ww}(S_u, D_{PLRS-PAR}) + \\
+ 3 \times T_{ww}(S_u, D_{NAR-PAR})
\end{aligned}
$$

Finally, the average latency to restore the downstream during a session can be expressed as follows:

$$
Avg\_Lat\_Down(MIP) = \frac{\sum_{1}^{N_m} Lat\_Down_i(MIP)}{N_m}
$$

$$
\begin{aligned}
Avg\_Lat\_Down(HMIP) = p_{\text{int}ra} \times \frac{\sum_{1}^{N_{\text{int}ra}} Lat\_Down_i(HMIP_{\text{int}ra})}{N_{\text{int}ra}} + \\
+ (1 - p_{\text{int}ra}) \times \frac{\sum_{1}^{N_{\text{int}er}} Lat\_Down_i(HMIP_{\text{int}er})}{N_{\text{int}er}}
\end{aligned}
$$

where p_{intra} is the intra-domain handover probability and $N_m = N_{intra} + N_{inter}$.

$$
\begin{aligned}
Avg\_Lat\_Down(FMIP) = p_{ant} \times \frac{\sum_{1}^{N_{ant}} Lat\_Down_i(FMIP_{ant})}{N_{ant}} + \\
+ (1 - p_{ant}) \times \frac{\sum_{Nreact}^{1} Lat\_Down_i(FMIP_{react})}{N_{react}}
\end{aligned}
$$

$$
\begin{aligned}
Avg\_Lat\_Down(GMP) = p_{ant} \times \frac{\sum_{i=1}^{N_{ant}} Lat\_Down_i(GMP_{ant})}{N_{ant}} \\
+ (1 - p_{ant}) \times \frac{\sum_{i=1}^{N_{react}} Lat\_Down_i(GMP_{react})}{N_{react}}
\end{aligned}
$$

where p_{ant} is the anticipated mode probability, $N_m = N_{ant} + N_{react}$, N_{ant} is the number of anticipated operations and N_{react} is the number of reactive operations.

5.3 Average handover latency to restore the upstream flow during a session

Average handover latency to restore the upstream flow (*Avg_Lat_Up*) during a session is defined as the average time of all handovers of a session to restore the MN's ability to send packets of the ongoing upstream. To reduce handover latency, some efforts were also made to enhance the L2-handover and the L3-handover. Here, we focus on L3-handover analysis and assume a constant value TL2 to express the L2-handover latency.

Once again, we assume that the routing optimization is used to reduce the latency of the packets sent/received by the MN to/from the CN. So, in order to restore the upstream, the MN should execute the return routability procedure before sending the BU message to the CN. This is the case of the MIP, $HMIP_{inter}$ and FMIP. So, the latency to restore the upstream of these protocols can be expressed as follows:

$$
\begin{aligned}
&Lat\_Up(MIP) = Lat\_Down(MIP) \\
&Lat\_Up(HMIP_{\text{int}er}) = Lat\_Down(HMIP_{\text{int}er}) \\
&Lat\_Up(FMIP_{ant}) = Lat\_Down(MIP) \\
&Lat\_Up(FMIP_{react}) = Lat\_Down(MIP)
\end{aligned}
$$

Because of the hierarchical characteristic of the $HMIP_{intra}$, the MN also does not need to execute the return routability procedure to restore the upstream. This stream is also restored just after the bi-directional tunnel is established between the MN and the MAP. So, the latency in this case can be expressed as follows:

$$Lat\_Up(HMIP_{\text{int}ra}) = Lat\_Down(HMIP_{\text{int}ra})$$

The upstream restore procedure is one of the main advantages of the GMP_{ant}. During the handover procedure, the network entities update the mobility binding cache of the AR of the CN (or CNs) on behalf of the MN. So, when the MN arrives at the new GAN, it can restore the upstream just after sending the *C2N_HandoverFinish* message and receiving the *N2C_HandoverAcknowledge* message. The latency in this case can be expressed as follows:

$$Lat\_Up(GMP_{ant}) = TL2 + 2 \times T_{wlw}(S_u, D_{MN-AR})$$

For the GMP_{react}, the network entities can only update the mobility binding cache of the AR of the CN (or CNs) after the MN has arrived at the new GAN. In order to restore the upstream, the following procedure is executed: (i) the MN sends the *C2N_HandoverFinish* message to the NAR. (ii) the NAR sends the *N2N_HandoverIndication* message to the NLRS to request the handover context of the MN at the PAR; (iii) the NLRS relays the *N2N_HandoverIndication* message to the PLRS; (iv) the PRLS relays de *N2N_HandoverIndication* message to the PAR; (v) the PAR sends the *N2N_Handover Acknowledge* message to the NAR to transfer the handover context of the MN; (vi) the NAR sends the *N2N_BindingNotification* message to the NLRS; (vii) the NLRS sends the *N2N_BindingNotification* message to the LRS of the CN; (viii) the LRS of the CN sends the *N2N_BindingNotification* message to the AR of the CN to update the binding of the MN; and (ix) the NAR sends the *N2N_HandoverAcknowledge* message to the MN. So, the latency in this case can be expressed as follows:

$$\begin{aligned} Lat\_Up(GMP_{react}) = {} & TL2 + 2\times T_{wlw}(S_u, D_{MN-NAR}) + \\ & + 2\times T_{ww}(S_u, D_{NAR-NLRS}) + \\ & + T_{ww}(S_u, D_{NLRS-PLRS}) + T_{ww}(S_u, D_{PLRS-PAR}) + \\ & + T_{ww}(S_u, D_{PAR-NAR}) + T_{ww}(S_u, D_{NLRS-CN\_LRS}) + \\ & + T_{ww}(S_u, D_{CN\_LRS-CN\_AR}) \end{aligned}$$

Finally, the average latency to restore the upstream during a session can be expressed as follows:

$$Avg\_Lat\_UP(MIP) = Avg\_Lat\_Down(MIP)$$

$$Avg\_Lat\_Up(HMIP) = Avg\_Lat\_Down(HMIP)$$

$$Avg\_Lat\_Up(FMIP) = Avg\_Lat\_Down(MIP)$$

$$\begin{aligned} Avg\_Lat\_UP(GMP) = {} & p_{ant} \times \frac{\sum_{i=1}^{N_{ant}} Lat\_Up_i(GMP_{ant})}{N_{ant}} \\ & + (1 - p_{ant}) \times \frac{\sum_{i=1}^{N_{react}} Lat\_Up_i(GMP_{react})}{N_{react}} \end{aligned}$$

where p_{ant} is the anticipated mode probability, $N_m = N_{ant} + N_{react}$, N_{ant} is the number of anticipated handovers and N_{react} is the number of reactive handovers.

5.4 Buffer size requirement for buffering and forwarding mechanisms

Some of the mobility protocols support buffering control of packets during handover time. This is the case of both the anticipated mode of operation of the FMIP and the GMP. These protocols use signaling messages to enable the buffering and forwarding of in-flight packets from the PAR to the NAR. We will denote the buffer size requirement for downstream (*Buf_Size_Down*) and upstream (*Buf_Size_Up*) as a function of the packet transmission rate, as follows.

For the anticipated mode of operation of the FMIP, both the PAR and the NAR should use buffers to store downstream in-flight packets. The PAR begins to store packets immediately after receiving the FBU message and remains buffering packets until the HAck message is received and the forwarding mechanism starts. In its turn, the NAR begins to store packets when the first in-flight downstream packet (forwarded by the PAR) is received and remains buffering packets until the FNA message is received and the delivering of packets starts. In this case, the buffer size requirement can be expressed as follows:

$$\begin{aligned} Buf\_Size\_Down(FMIP) = {} & [T_{wlw}(S_u, D_{MN-PAR}) + \\ & + 2\times T_{ww}(S_u, D_{NAR-PAR})]\times R_d + \\ & + [TL2 + T_{wlw}(S_u, D_{MN-NAR})]\times R_d \end{aligned}$$

The buffering mechanism of the anticipated mode of operation of the GMP also stores in-flight downstream packets at both the PAR and the NAR. The PAR begins to store packets immediately after receiving the *C2N_HandoverInitiate* message and remains buffering packets until the *N2N_HandoverAcknowledge* message is received, when the forwarding mechanism starts. In turn, the NAR begins to store packets when the first in-flight downstream packet (forwarded by the PAR) is received and remains buffering packets until the *C2N_HandoverFinish* message is received, when the delivering of the buffered packets starts. So, for the GMP, the buffer size requirement can be expressed as follows:

$$\begin{aligned} Buf\_SIze\_Down(GMP) = {} & [T_{wlw}(S_u, D_{MN-PAR}) + T_{wlw}(S_u, D_{PAR-PLRS}) + \\ & + T_{ww}(S_u, D_{PLRS-NLRS}) + T_{ww}(S_u, D_{NLRS-NAR}) + \\ & + 3\times T_{ww}(S_u, D_{NAR-PAR})]\times R_d + \\ & + [TL2 + T_{wlw}(S_u, D_{MN-NAR})]\times R_d \end{aligned}$$

In both the FMIP and the GMP, the buffer size requirement for the upstream flow should be supported by the MN. For the FMIP, the requirement is independent from the mode of operation (anticipated or reactive) and the upstream flow is restored after the return routability and binding update procedures. For the GMP, the buffer requirement depends on the mode of operation. Under the anticipated mode, the GMP starts the buffering mechanism after the L2-Down trigger is fired by the link layer and finishes after the NAR acknowledges the *C2N_HandoverFinish* message sent by the MN. This optimization is a result of the anticipated registering of the MN at the GAN and the notification of the ARs of the CNs during L2 handover. For the reactive mode, the buffering mechanism also starts after the L2-Down trigger but only finishes after the binding notification procedure finishes and the NAR acknowledges the *N2N_HandoverFinish* message sent by the MN. So, for these protocols, the buffer size requirement for the upstream can be expressed as follows:

$$\begin{aligned} Buf\_Size\_Up(FMIP) = [&TL2 + 2\times T_{wlw}(S_u, D_{MN-AR}) + \\ &+ (2+3+3+2)\times T_{wlw}(S_u, D_{MN-HA}) + \\ &+ 2\times T_{ww}(S_u, D_{HA-CN}) + \\ &+ (2+2)\times T_{wlw}(S_u, D_{MN-CN})]\times R_u \end{aligned}$$

$$Buf\_Size\_Up(GMP_{ant}) = [TL2 + 2\times T_{wlw}(S_u, D_{MN-AR})]\times R_u$$

$$\begin{aligned} Buf\_Size\_Up(GMP_{react}) = [&TL2 + 2\times T_{wlw}(S_u, D_{MN-NAR}) + \\ &+ 2\times T_{ww}(S_u, D_{NAR-NLRS}) + T_{ww}(S_u, D_{NLRS-PLRS}) + \\ &+ T_{ww}(S_u, D_{PLRS-PAR}) + T_{ww}(S_u, D_{PAR-NAR}) \\ &+ T_{ww}(S_u, D_{NLRS-CN\_LRS}) + T_{ww}(S_u, D_{CN\_LRS-CN\_AR})]\times R_u \end{aligned}$$

5.5 Result Analysis

Our experiment uses the equations of the performance parameters presented in sections 5.1 , 5.2 , 5.3 and 5.4 to analyze the mobility scenario illustrated in Figure 18. In this scenario, the MN moves from the access network of the AR11 towards the access network of the AR43. We assume that the MN establishes a bidirectional communication session with the stationary CN. This communication session begins at the network of the AR11 and terminates when the MN arrives at the network of the AR43. We also assume that the average movement speed of the

MN is adequate to allow the whole execution of the handover process between each access network.

For each of the mobility protocols analyzed in our work, we used the equations to calculate the time to execute the intra-domain and inter-domain handover process. Along the path covered by the MN, eleven handovers takes place, eight of which are intra-domain and three others are inter-domain. Our analysis of the handover process calculates the time to execute three parts of the handover: (i) the time to start the buffering and forwarding mechanisms; (ii) the time to restore the downstream flow of packets; and (iii) the time to restore the upstream flow of packets.

The time values calculated for each part of the handover of the same mobility protocol considerably varies according to the location the handover takes place because the signaling messages exchanged between the entities that controls the mobility of the MN must have to go through short or long paths to reach their destiny. So, for the first approach of our analysis, we calculated the time average values to execute each of the three parts of the handover process, for each mobility protocol, along the path covered by the MN. The values of the parameters used in the equations are listed in Table 6 and the D_{x-y} values (average number of hops between x and y) used are illustrated in Figure 19. The average time values calculated are listed in Table 7.

The value used for the L2 handover time (*TL2*) is the sum of the average time of the search and execution phases of the L2 handover process of four different commercial IEEE 802.11b cards with different chipsets, as described in [15]. The detection phase is discarded because we assume that the MIES is available. So, this phase is substituted by the L2-GoingDown trigger. The L2-Down and L2-UP triggers are also available.

Table 6 – Parameters settings used in the first approach of the analysis

Parameter	Value
S_u	48 bytes
B_w	100 Mbps
L_w	0.5 msec
B_{wl}	11 Mbps
L_{wl}	2 msec
P_t	0,001 msec
$TL2$	186 msec

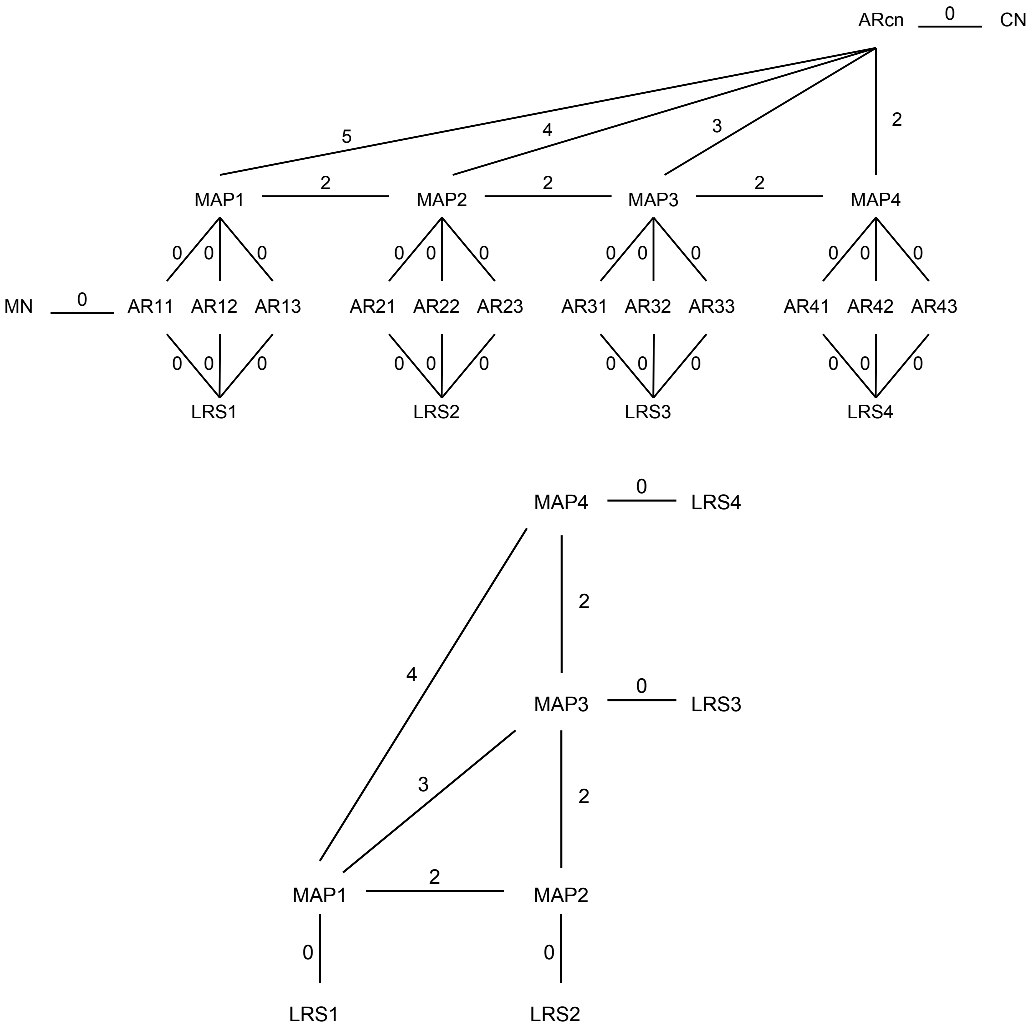

Figure 19 – Dx-y values used in the analysis

The average time values of Table 7 demonstrate that the time to start the buffering and forwarding mechanisms in the FMIP and GMP is quite similar in both modes of operation. This indicates (and will be further demonstrated) that the loss of downstream packets will be very similar. The results also demonstrate that the time to restore the downstream packets in the HMIP during inter-domain handover is greater than the MIP. This is expected because HMIP uses more signaling messages than the MIP to optimize the intra-domain handover, in which case the HMIP shows a better performance than the MIP.

Table 7 – Average time values calculated to the three parts of the handover process

	Average time values calculated (ms)					
	Start buffering/ forwarding mechanism		**Restore downstream flow of packets**		**Restore upstream flow of packets**	
	Intra	**Inter**	**Intra**	**Inter**	**Intra**	**Inter**
MIP	NA	NA	227	231	259	263
HMIP	NA	NA	210	236	242	268
$\text{FMIP}_{\text{anticipated}}$	0	0	188	188	258	258
$\text{FMIP}_{\text{reactive}}$	190	193	190	193	260	263
$\text{GMP}_{\text{anticipated}}$	0	0	188	188	188	188
$\text{GMP}_{\text{reactive}}$	193	199	193	199	206	213

Just after the L2 handover, both the FMIP and GMP are ready to restore the downstream flow. So, the average time of downstream reestablishment is almost the same as the time needed to complete L2 handover. Their performances are better than the MIP and HMIP because they don’t need to execute the return routability procedure to restore the downstream flow. On the other hand, FMIP is required to execute this procedure to restore the upstream flow the same way they do. Therefore, the FMIP upstream-flow restoration time is quite similar to the time calculated for the MIP and the HMIP. Here, we call attention to the first clear advantage of the GMP. When operating in the anticipated mode, the registration and binding notifications are executed by the NAR and the NLRS of the new GAN on behalf of the MN (during the L2 handover process), which enables the MN to restore the upstream flow of packets immediately after the L2 handover. Even when operating in the reactive mode, this time is lower than the ones in the other protocols, which indicates that the binding notification procedure of the GMP performs better than the binding update procedure.

The second point of our analysis calculates the loss of packets of the downstream flow and the time required to restore the downstream and upstream flows for different transmission rate conditions. We also analyze the performance of these protocols for different probability occurrences of the intra-domain and inter-domain handover, as well as of the anticipated and reactive mode of operations.

We assume that the bidirectional communication session between the MN and the CN takes 1000 seconds. The residence time of the MN in each access network is 20 seconds. Therefore, the number of movements between access networks during a communication session is 50. Each L3 handover between access networks is executed as an intra-domain handover with probability P_{intra} and as an inter-domain handover with probability $(1 - P_{intra})$. In the same manner, when relevant, the anticipated mode of operation occurs with probability $P_{antecipated}$ while the reactive mode of operation occurs with probability $(1 - P_{antecipated})$.

Table 8 – Parameters settings used in the second analysis

Parameter	Value
T_s	1000 sec
T_r	20 sec
N_m	50
R_d	8 – 1024 kbps
R_u	8 – 1024 kbps
P_{intra}	0 – 1 (default 0.8)
$P_{anticipated}$	0 – 1 (default 0.8)

The values of the parameters used in the equations for this second analysis are the same used for the first one plus the average time values listed in Table 7 and the parameters listed in Table 8.

Figure 20 shows the comparison of the total packet loss of the downstream flow for a communication session. Because of the absence of the buffering and forwarding mechanisms, MIP and HMIP loose a significant amount of packets if compared to FMIP and GMP when the downstream transmission rate is above 64 Kbps. For the FMIP and GMP, the loss of packets is equivalent because their buffering and forwarding mechanisms are essentially the same. These protocols loose packets from the downstream flow during handover only in the reactive mode. In the anticipated mode, we assume that the amount of buffers required to store packets are always available and none of the forwarding packets are lost. This analysis used the following probabilities: $P_{intra} = 0.8$ and $P_{antecipated} = 0.8$.

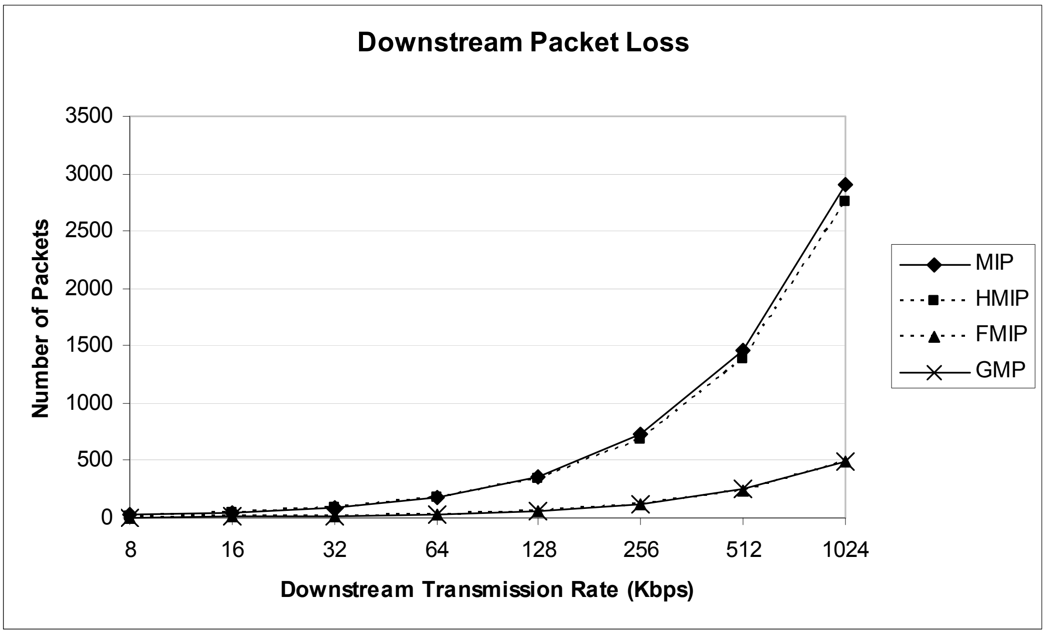

Figure 20 – Downstream flow packet loss versus transmission rate

As shown in Figure 21, the average buffer size requirement of the FMIP to store packets of the upstream flow is significantly bigger than the one required for the GMP. The FMIP requires about 38% more buffer area to store these packets because, after the L2 handover, it is required to execute the return routability procedure to notify the CN about the changing of the binding. So, it must continue the buffering mechanism for an extra period until the procedure is finished. On the other hand, the GMP starts the forwarding mechanism immediately after the L2 handover, when the anticipated mode of operation is in progress.

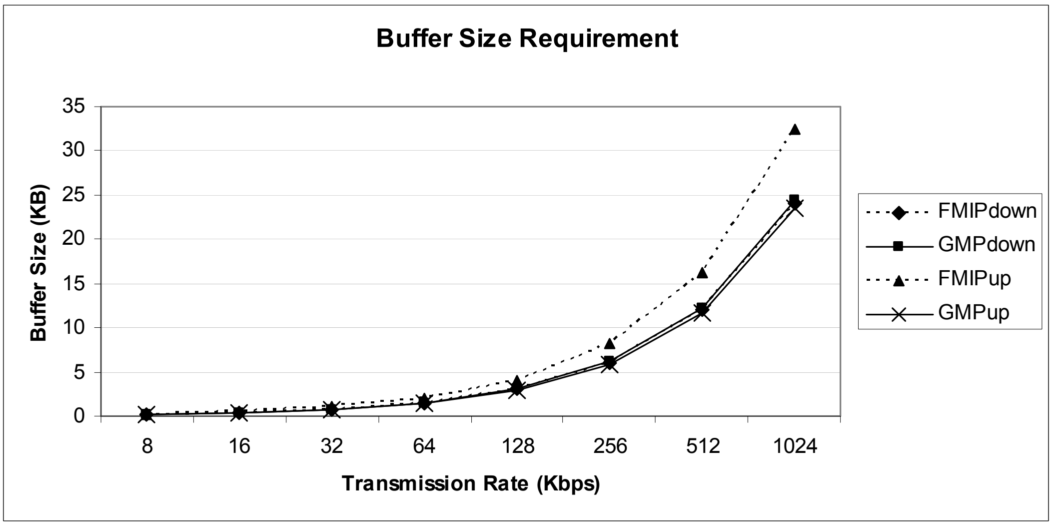

Figure 21 – Buffer size requirement for upstream/downstream flow

Figure 22 shows the comparison of the latency to restore the downstream flow based on the probability of intra-domain handover and the operating probability of the anticipated mode. This analysis demonstrates that most of the time the HMIP performs better than the MIP. The hierarchical functionality of the HMIP restores the communication faster than the MIP because, most of the time, the registering of the MN and the establishment of the bidirectional tunnel is local to the access network. Unlike the HMIP, the MIP always registers the MN and establishes the bidirectional tunnel with the HA. FMIP and GMP perform significantly better than both the other two. Basically, those protocols make use of the buffering and forwarding mechanisms to reduce the latency. The performance of the FMIP is slightly better than the performance of the GMP in the reactive mode of operation. Notice that the GMP defines the protocol messages to use the service of the LRS to discover the address of the ARs based on the PoA-ID and to notify the PAR about the handover process. These messages take part in the handover process and are time consuming. Although these tasks must also be executed by the FMIP, there are no entities defined to execute them. So, they are implicitly executed by the AR itself which increases the complexity of this entity. Because of this overburden of the AR, the FMIP has fewer signaling messages than the GMP which leads to a lower latency.

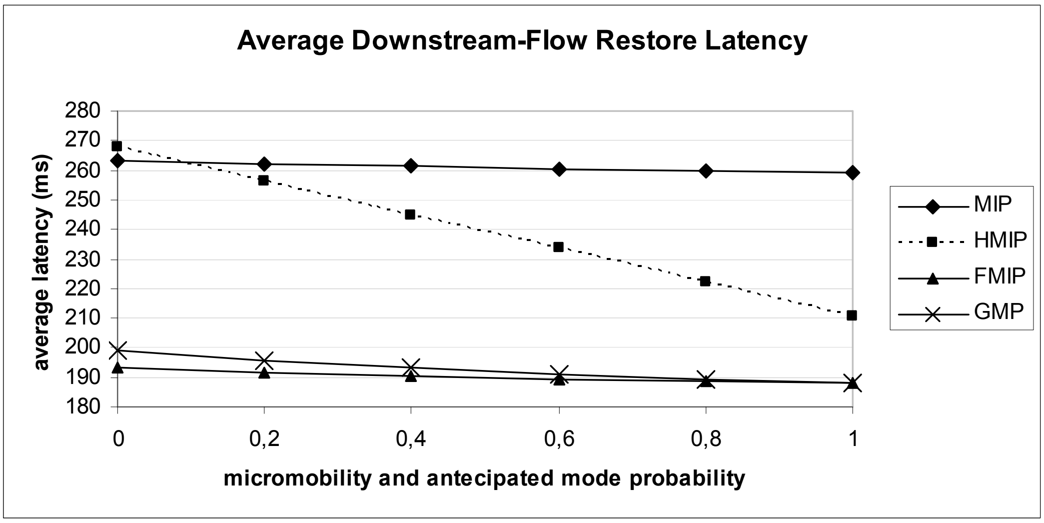

Figure 22 – Average downstream-flow restore latency

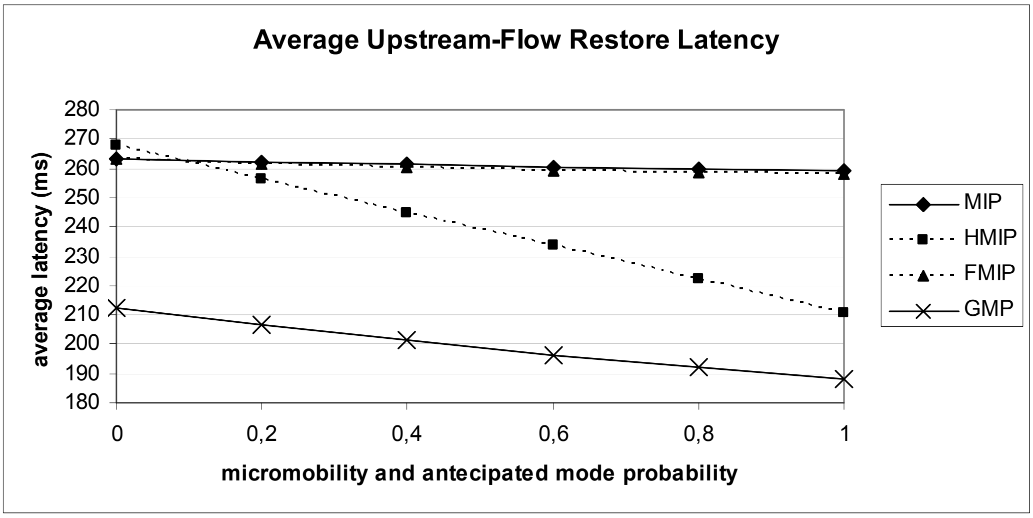

Figure 23 - Average upstream-flow restore latency

In the following, we compare the latency to restore the upstream flow based on the probability of the intra-domain handover and the probability of the anticipated mode. This analysis demonstrates that the MIP and the FMIP are equivalent in this case. Here, the FMIP is required to register the MN at the HA and to establish the bidirectional tunnel with the HA in order to execute the return routability procedure. These procedures are time consuming and significantly impact the performance of the FMIP. So, now, the HMIP reaches a better average performance than the MIP and FMIP. Here, the advantage of the anticipated registering and binding notification procedures executed by the ARs and the LRSs on behalf of the MN makes the difference in the performance of the GMP.